



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Scottish Gaelic Speech and Writing

Citation for published version:

Lamb, W 2008, *Scottish Gaelic Speech and Writing: Register Variation in an Endangered Language*. Cló Ollscoil na Banríona, Belfast.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Scottish Gaelic Speech and Writing: Register Variation in an Endangered Language

William Lamb



Belfast Studies in Language, Culture and Politics
General Editors: John M. Kirk and Dónall P. Ó Baoill

- 1: *Language and Politics: Northern Ireland, the Republic of Ireland, and Scotland* published 2000 ISBN 0 85389 791 3
- 2: *Language Links: the Languages of Scotland and Ireland* published 2001 ISBN 0 85389 795 6
- 3: *Linguistic Politics: Language Policies for Northern Ireland, the Republic of Ireland, and Scotland* published 2001 ISBN 0 85389 815 4
- 4: *Travellers and their Language* published 2002 ISBN 0 85389 832 4
- 5: Simone Zwickl, *Language Attitudes, Ethnic Identity and Dialect Use across the Northern Ireland Border: Armagh and Monaghan* published 2002 ISBN 0 85389 834 0
- 6: *Language Planning and Education: Linguistic Issues in Northern Ireland, the Republic of Ireland, and Scotland* published 2002 ISBN 0 85389 835 9
- 7: Edna Longley, Eamonn Hughes and Des O'Riordan (eds.) *Ireland (Ulster) Scotland: Concepts, Contexts, Comparisons* published 2003 ISBN 0 85389 844 8
- 8: Maolcholaim Scott and Roise Ní Bhaoill (eds.) *Gaelic-Medium Education Provision: Northern Ireland, the Republic of Ireland, Scotland and the Isle of Man* published 2003 ISBN 0 85389 847 2
- 9: Dónall Ó Riagáin (ed.) *Language and Law in Northern Ireland* published 2003 ISBN 0 85389 848 0
- 10: *Towards our Goals in Broadcasting, the Press, the Performing Arts and the Economy: Minority Languages in Northern Ireland, the Republic of Ireland, and Scotland* published 2003 ISBN 0 85389 856 1
- 11: J. Derrick McClure (ed.) *Doonsin' Emeralds: New Scribes anent Scots an Gaelic / New Studies on Scots and Gaelic* published 2004 ISBN 0 85389 860 X
- 12: Neal Alexander, Shane Murphy and Anne Oakman (eds.) *To the Other Shore: Cross-currents in Irish and Scottish Studies* published 2004 ISBN 0 85389 863 4
- 13: *Legislation, Literature and Sociolinguistics: Northern Ireland, the Republic of Ireland, and Scotland* published 2005 ISBN 0 85389 874 X
- 14: Shane Alcobia-Murphy, Johanna Archbold, John Gibney and Carole Jones (eds.) *Beyond the Anchoring Grounds: More Cross-currents in Irish and Scottish Studies* published 2005 ISBN 0 85389 885 5
- 15: Dónall Ó Riagáin (ed.) *Voces Diversae: Lesser-Used Language Education in Europe* published 2006 ISBN 0 85389 886 3
- 16: William Lamb *Scottish Gaelic Speech and Writing: Register Variation in an Endangered Language* published 2008 ISBN 978 0 85389 895 5
- 17: Alasdair MacCaluim *Reversing Language Shift: The Role and Social Identity of Scottish Gaelic Learners* published 2007 ISBN 978 0 85389 897 9
- 18: Seán Mac Corraigh *Ar thoir an dea-chleachtais: The quest for Best Practice in Irish-medium Primary Schools in Belfast* published 2008 ISBN 978 0 85389 898 6
- 19: *Language and Economic Development: Northern Ireland, the Republic of Ireland, and Scotland* published 2008 ISBN 978 0 85389 910 X

Scottish Gaelic Speech and Writing: Register Variation in an Endangered Language

William Lamb
University of the Highlands and Islands

Cló Ollscoil na Banríona
2008

First published in 2008
Cló Ollscoil na Banríona
Queen's University Belfast
Belfast, BT7 1NN

Belfast Studies in Language, Culture and Politics
www.bslcp.com

© Cló Ollscoil na Banríona and William Lamb

The publication of this volume has been made possible through the financial support of Bòrd na Gàidhlig.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN 978 0 85389 895 5

Typeset by William Lamb and Nigel Craig in Granjon
Cover artwork by Kate MacDonald
Cover design by Colin Young
Printing by the Cromwell Press, Trowbridge, Wiltshire

Table of Contents

LIST OF TABLES	9
LIST OF FIGURES	
ABBREVIATIONS	11
ACKNOWLEDGEMENTS	14
AUTHOR'S NOTE	15
PREFACE BY JIM MILLER	16
1 INTRODUCTION	17
1.1 BACKGROUND AND RATIONALE	17
1.2 RESEARCH QUESTIONS	21
1.3 ASSUMPTIONS AND LIMITATIONS	22
1.4 DEFINITIONS	22
2 SPOKEN AND WRITTEN REGISTERS AND SCOTTISH GAELIC	23
2.1 SPEECH AND WRITING	23
2.1.1 <i>Linguistic Differences between Writing and Speech: Empirical Studies</i>	24
2.1.2 <i>Complexity and Subordination in Speech and Writing</i>	28
2.1.3 <i>Spontaneous Spoken Language (SSL)</i>	29
2.2 REGISTER: CONTEXT-BASED LINGUISTIC VARIATION	31
2.3 MULTI-DIMENSIONAL (MD) ANALYSIS	32
2.4 SCOTTISH GAELIC: PREVIOUS REGISTER AND SPOKEN LANGUAGE RESEARCH	35
2.5 SUMMARY	37
3 SOCIOLINGUISTICS	39
3.1 BRIEF HISTORY AND DIALECTOLOGY	39
3.2 DEMOGRAPHY AND ETHNOGRAPHY	42
3.3 ECONOMY AND COMMUNITY	42
3.4 LANGUAGE USAGE	44
3.4.1 <i>Age, Occupation, and Gender</i>	44
3.4.2 <i>Register Range and Bilingualism</i>	47
3.5 SUMMARY	50
4 METHODOLOGY	52
4.1 DATA COLLECTION AND COMPOSITION	52
4.1.1 <i>Origin and Selection of Texts</i>	53
4.1.2 <i>Transcription Practices</i>	55
4.1.3 <i>Situational Description of the Registers with Examples of Texts</i>	55
4.2 SELECTION OF LINGUISTIC FEATURES	63
4.3 CORPUS MARK-UP	66
4.3.1 <i>Issues and Limitations Posed by Mark-up</i>	68
4.3.2 <i>Grammatical Decisions: Adverbs and Case</i>	68
4.4 TAG COUNTING	69
4.5 DATA ANALYSIS	70
INTRODUCTION TO THE RESULTS CHAPTERS	71
5 INFORMATION STRUCTURE AND CLAUSAL TYPES	73
5.1 CLEFTS	73
5.2 DETACHMENT	77
5.3 PLACEMENT OF ADVERBIALS	81
5.4 CLAUSE TYPES	84

5.4.1	<i>Coordinate Clauses</i>	84
5.4.2	<i>Cosubordinate Clauses</i>	87
5.4.3	<i>Main/matrix Clauses</i>	88
5.4.4	<i>Subordinate Clauses</i>	90
5.4.5	<i>Juxtaposition versus Subordination in Spoken Scottish Gaelic</i>	97
5.5	SUMMARY OF CHAPTER FIVE	100
6	MORPHOSYNTAX AND THE LEXICON	102
6.1	VERBAL MORPHOSYNTAX	102
6.1.1	<i>Aspect</i>	102
6.1.2	<i>Tense</i>	105
6.1.3	<i>Valence Decreasing Expressions</i>	107
6.1.4	<i>Clausal Negation</i>	111
6.1.5	<i>Mode: Interrogative clauses and Imperative Verbs</i>	112
6.1.6	<i>Copular Verbs</i>	114
6.1.7	<i>Identificational Clefts</i>	115
6.2	NOMINAL MORPHOSYNTAX	116
6.2.1	<i>Adjectives</i>	116
6.2.2	<i>Appositives</i>	120
6.2.3	<i>Complex Noun Phrases</i>	122
6.2.4	<i>Nouns: Total Nouns, Nonconcordant Nouns, Genitives and Indefinite Marking</i>	123
6.3	POSSESSION, PREPOSITIONS, AND PRONOUNS	128
6.3.1	<i>Possession</i>	128
6.3.2	<i>Prepositions</i>	131
6.3.3	<i>Pronouns</i>	132
6.4	ADVERBS	135
6.5	LEXICAL CLASSES	137
6.5.1	<i>Adverbs of Stance and Degree: Amplifiers, Downtoners, Emphatics, and Hedges</i>	137
6.5.2	<i>Verb Types</i>	140
6.5.3	<i>English Borrowing</i>	143
6.5.4	<i>Numerals</i>	145
6.6	LEXICAL SPECIFICITY	146
6.7	SUMMARY OF CHAPTER SIX	148
7	NOUN PHRASE GRAMMAR AND COMPLEXITY	150
7.1	CASE MARKING IN SCOTTISH GAELIC REGISTERS: CONCORDANCE AND NONCONCORDANCE	150
7.1.1	<i>Nominative Case</i>	151
7.1.2	<i>Dative Case</i>	152
7.1.3	<i>Genitive Case</i>	153
7.1.4	<i>Frequency of Nonconcordance per Tag Type</i>	156
7.1.5	<i>Adjectival Case Marking</i>	156
7.1.6	<i>Summary of Noun Phrase Grammar</i>	157
7.2	NOUN PHRASE COMPLEXITY AND FUNCTION	157
7.2.1	<i>Data Coding</i>	157
7.2.2	<i>Tabular Data</i>	158
7.2.3	<i>Examples from the Text</i>	159
7.2.4	<i>Summary of NP Complexity and Function</i>	171
8	PROFILES OF THE INDIVIDUAL REGISTERS	172
8.1	CONVERSATION	172
8.2	RADIO INTERVIEW	175
8.3	SPORTS REPORTAGE	176
8.4	TRADITIONAL NARRATIVE	179
8.5	FICTION	180
8.6	FORMAL PROSE	181
8.7	NEWS SCRIPTS	182
8.8	POPULAR WRITING	184

9	CONCLUSIONS	186
9.1	SUMMARY	186
9.2	ASSOCIATIONS BETWEEN CONTEXT AND LINGUISTIC FORM	188
9.3	A REVIEW OF THE FINDINGS IN TERMS OF PAST REGISTER RESEARCH	190
9.4	CONCLUDING REMARKS ON SPONTANEOUS SPOKEN GAELIC	194
9.5	IMPLICATIONS AND SUGGESTIONS FOR FUTURE STUDY	195
	APPENDIX 1: A DESCRIPTIVE GRAMMAR OF SCOTTISH GAELIC	197
1	PHONOLOGY: A BRIEF SUMMARY	198
1.1	VOWELS	198
1.2	CONSONANTS	199
1.3	DIPHTHONGS	199
1.4	WORD STRESS	199
1.5	CONSONANT MUTATION	199
1.6	PREASPIRATION	201
2	MORPHOLOGY	202
2.1	OVERVIEW OF MORPHOLOGICAL CHARACTERISTICS	202
2.1.1	<i>Synthesis and Fusion</i>	202
2.1.2	<i>Morphological Processes</i>	202
2.1.3	<i>Head/Dependent Marking</i>	204
2.2	NOMINAL MORPHOLOGY	205
2.2.1	<i>Nouns</i>	205
2.2.2	<i>Pronouns</i>	214
2.2.3	<i>Numerals</i>	218
2.2.4	<i>Adjectives</i>	219
2.2.5	<i>Adverbs</i>	222
2.2.6	<i>Prepositions</i>	224
2.2.7	<i>Nominal Categorisors: Nominalisation and Diminutives</i>	228
2.3	VERBAL MORPHOLOGY	228
2.3.1	<i>Verbs</i>	229
2.3.2	<i>Verbal Categorisors</i>	241
2.3.3	<i>Verbal Modifiers: Participles and Gerunds</i>	244
3	SYNTAX	245
3.1	PREDICATE NOMINALS AND SIMILAR CONSTRUCTIONS	245
3.1.1	<i>Predicate Nominals: Proper Inclusion and Equation</i>	245
3.1.2	<i>Predicate Adjectives</i>	246
3.1.3	<i>Locatives and Possessive Predication</i>	246
3.1.4	<i>Existentials</i>	247
3.2	CONSTITUENT ORDER AND GRAMMATICAL RELATIONS	248
3.2.1	<i>Postposition of Weak Pronominal Objects</i>	249
3.2.2	<i>Small Clauses</i>	250
3.2.3	<i>'Composite Verbs'</i>	251
3.3	CLAUSE COMBINATIONS	251
3.3.1	<i>Coordination</i>	254
3.3.2	<i>Subordination</i>	256
3.3.3	<i>Cosubordination</i>	263
3.4	DISCOURSE PHENOMENA	264
3.4.1	<i>Referential Devices</i>	264
3.4.2	<i>Highlighting of Constituents</i>	266
3.4.3	<i>Discourse Particles, Parenthetical Statements, and Interjections</i>	271
3.5	INFLUENCE OF ENGLISH	273
4	SAMPLE TEXT: ORAL FOLK TALE	275
	SUPPLETIVE ('IRREGULAR') VERBS	279
	A. ACTIVE VOICE	279
	B. PASSIVE VOICE: IMPERSONAL FORMS	280

THE 100 MOST COMMON WORDS OF SCOTTISH GAELIC	281
INDEX OF GAELIC WORDS USED IN GRAMMAR	283
APPENDIX 2: LIST OF TEXTS IN CORPUS	287
APPENDIX 3: FULL TAG SET	291
REFERENCES	296
GENERAL INDEX	305

List of Tables

1 Comparison of text types in Jang (1998) and Biber (1995)	34
2 Language-Maintenance and Viability Survey 1986–1988	45
3 Euromosaic Survey 1994–1995	45
4 Organisation and Make-up of the Scottish Gaelic Corpus	53
5 Corpus tags which were used for the quantitative analysis	64
6 Example of a tagged text	67
7 Results: clefts	74
8 Sorted data: 's ann clefts	74
9 Sorted data: demonstrative clefts	75
10 Sorted data: 'e clefts	75
11 Results: detachment	78
12 Sorted data: left-detachment	78
13 Results: fronted adverbials and peripheral modifiers	81
14 Sorted data: fronted adverbials and peripheral modifiers	81
15 Results: placement of subordinate clauses	82
16 Results: coordinate clauses	85
17 Sorted data: coordinate clauses	85
18 Results: cosubordinate clauses	87
19 Results: main/matrix clauses	88
20 Sorted data: main/matrix clauses	89
21 Results: temporal subordination	90
22 Sorted data: temporal subordination	91
23 Results: relative clauses	93
24 Sorted data: relative clauses	93
25 Results: complement clauses	94
26 Sorted data: complement clauses	94
27 Results: concessive clauses	95
28 Results: conditional clauses	96
29 Results: cause and reason clauses	97
30 Results: aspect	102
31 Sorted data: perfect aspect	103
32 Sorted data: progressive aspect	104
33 Results: tense	105
34 Sorted data: present tense	106
35 Sorted data: past tense	106
36 Sorted data: indefinite 2	107
37 Results: valence decreasing constructions	108
38 Sorted data: passive clauses with agents	109
39 Sorted data: total passive clauses	109
40 Sorted data: synthetic passives and impersonals	110
41 Results: clausal negation	111
42 Sorted data: clausal negation	111
43 Results: interrogatives and imperatives	112
44 Sorted data: interrogative clauses	112
45 Sorted data: imperative clauses	114
46 Results: copular verbs	115
47 Sorted data: copular verbs	115
48 Results: identificational clefts	116

49 Results: adjectives	117
50 Sorted data: attributive adjectives	117
51 Sorted data: demonstrative adjectives	117
52 Sorted data: predicative adjectives	118
53 Proportion of demonstrative adjectives in formal prose and conversation	119
54 Results: appositives	121
55 Results: complex noun phrases	122
56 Sorted data: complex noun phrases	123
57 Results: total nouns	123
58 Sorted data: total nouns	124
59 Results: grammatical nonconcordance	125
60 Sorted data: grammatical nonconcordance	126
61 Results: genitive nouns	126
62 Sorted data: genitive nouns	127
63 Results: indefinite marking	127
64 Sorted data: indefinite marking	128
65 Results: possession	129
66 Sorted data: aig possession	129
67 Sorted data: simple possession	129
68 Results: prepositions	131
69 Sorted data: prepositions	131
70 Results: verbal nouns per 1000 words	132
71 Results: pronouns	133
72 Sorted data: demonstrative pronouns	133
73 Sorted data: total pronouns	134
74 Results: adverbs	135
75 Sorted data: place adverbs	136
76 Sorted data: time adverbs	137
77 Results: adverbs of stance and degree	138
78 Sorted data: amplifiers	138
79 Sorted data: hedges	139
80 Results: verb types	141
81 Sorted data: private verbs	141
82 Sorted data: public verbs	142
83 Results: English borrowing	143
84 Assimilated English nouns	144
85 Unassimilated English nouns	144
86 Unassimilated English non-nouns	145
87 Results: numerals	146
88 Sorted data: numerals	146
89 Results: lexical specificity	147
90 Sorted data: average characters per word	147
91 Sorted data: type/token ratio	147
92 Noun concordance data: nominative case, feminine	151
93 Noun concordance data: nominative case, masculine	151
94 Noun concordance data: dative case, feminine	152
95 Noun concordance data: dative case, masculine	153
96 Noun concordance data: genitive case, feminine	154
97 Noun concordance data: genitive case, masculine	154
98 Noun concordance data: feminine, singular, definite, genitive	154
99 Noun concordance data: feminine, singular, indefinite, genitive	155

100 Noun concordance data: masculine, singular, definite, genitive	155
101 Noun concordance data: masculine, singular, indefinite, genitive	155
102 Average C: N ratio for genitive nouns	155
103 Nonconcordance per tag type	156
104 Adjectival grammatical concordance	156
105 Constituent frequencies for complex NPs	158
106 Syntactic functions of complex NPs	159
107 Single adjective noun phrases from spontaneous speech	160
108 Single adjective noun phrases from formal prose	161
109 Noun phrases modified by a prepositional phrase in spontaneous speech	162
110 Noun phrases modified by a prepositional phrase in formal prose	163
111 NPs modified by a relative clause in spontaneous speech	165
112 NPs modified by a relative clause in formal prose	165
113 Multiple genitives in formal prose	166
114 NPs modified by two or more adjectives in spontaneous speech	167
115 NPs modified by two or more adjectives in formal prose	168
116 NPs modified by two or more prepositional phrases in formal prose	170
117 NPs modified by two or more relative clauses in spontaneous speech	170
118 NPs modified by two or more relative clauses in formal prose	171
119 Biber's (1988) Dimension One: Informational vs. Involved Production	191
120 Biber's (1988) Dimension Two: Narrative vs. Non-narrative Concerns	192
121 Biber's (1988) Dimension Three: Explicit vs. Situation Dependent Reference	192
122 Consonant inventory	199
123 Lenition of Initial Consonants	200
124 Vowel Changes in Palatalisation (based in part on Ó Maolalaigh 1996)	201
125 Preaspiration after back vowels	201
126 The article: masculine noun, consonant-initial	208
127 The article: masculine noun, vowel-initial	209
128 The article: masculine noun, /f/-initial	209
129 The article: feminine noun, consonant-initial	209
130 The article: feminine noun, vowel-initial	209
131 The article: feminine noun, /f/-initial	209
132 Possession	213
133 Personal PN and emphatic suffix	214
134 Demonstratives	215
135 Prepositional pronouns	226
136 Possessive prepositional PNs	227
137 Pre-verbal particles	230
138 Weak verb inflection	234
139 Relativisation strategies	263

List of Figures

Figure 1: Tag Counting in LinguaStat	69
Figure 2: KWIC-type Concordances in LinguaStat	69
Figure 3: Tree diagram of the Scottish Gaelic Clause	252

List of Maps

Map 1: Traditional Gaelic-speaking Scotland	39
Map 2: The Southern Isles of the Outer Hebrides	41

List of Abbreviations

Where abbreviations in the running text were felt to be in common currency or clear from the surrounding context, they are not listed here. Abbreviations that are in small-caps in the glosses are often in small letters in the text, followed by a full-stop (e.g. 'adj.' for 'adjective').

*	utterance not grammatical	CUP	Cambridge University Press
:	incomplete word	DAT, D	dative (prepositional) case
?	questionable utterance;	DEF	definite
	interrogative clause	DEM	demonstrative
§	section (non-appendix)	DEP	dependent
1	first person	DERIV	derivational suffix
1P	first person plural	DIM	diminutive
1S	first person singular	DIR	directional
2	second person	DO	direct object
2P	second person plural	EMPH	emphatic suffix
2S	second person singular	EXIST	existential
3P	third person plural	FEM, F	feminine
3S	third person singular	FICT	fiction
A	most agentive argument of a	FOC	focus
	multi-argument clause	FORM	formal prose
A§	section of Appendix 1	GEN, G	genitive
ACC	accusative case	IMP	imperative
ADJ	adjective	IMPERS	impersonal
ADV	adverb	INDEP	independent
AGR	agreement	INDEF1	first indefinite ('present- future')
ARG	argument	INDEF2	second indefinite (<i>'conditional/habitual'</i>)
ART	article	INF	infinitive
ASP	aspect	IO	indirect object
AUX	auxiliary	INTER	radio interview
CLM	clause linkage marker	^L	lenition
COMP	complementiser	Lit	literally
CONJ	conjunction	LD	left-detachment
CONV	conversation		
COP	copular verb		

LOC	locative	PROG	progressive aspect
MASC, M	masculine	QUAL	quality
MKR	marker	QP	question particle
N	noun	RECIP	reciprocal
NEG	negative	REFL	reflexive
NEWS	news scripts	REL	relativiser
NOM, N	nominative case	S	only argument of an intransitive clause
NP	noun phrase	S	singular
n.s.	non significant (statistically)	SA	School of Scottish Studies archive tape
NUC	nucleus	ScG	Scottish Gaelic
OBJ	object	SPORT	sports reportage
OBL	oblique	SSL	spontaneous spoken language
P	least agentive argument of a multi-argument clause	STD	standard
PART	past participle	SUB	subject
PAST	definite past tense	SUPER	superlative
PERF	perfect aspect	TAM	tense, aspect, mode
PL	plural	TRAD	traditional narrative
PN	pronoun	V	verb
POP	popular writing	VN	verbal noun
POSS	possessive	VOC, V	vocative case
PP	prepositional phrase	WH	'what' as in WH-cleft
PRED	predicate		
PREP	preposition		
PRES	definite present tense		

Acknowledgements

For the present book, I am indebted to John Kirk, of Cló Ollscoil na Banríona, who was the prime mover in organising the funding necessary to publish the manuscript, and who also provided many helpful comments for getting it into shape. I am equally indebted to Wilson McLeod for acting as an unofficial editor and for commenting in very careful detail and so beneficially on successive sets of proofs throughout what has been a long production period. Many thanks are due as well to Jim Miller for his preface, Fiona Barr for producing the final index and Bòrd na Gàidhlig for providing the necessary publication funding.

I am still indebted to a number of friends and colleagues for their help during the initial research period and writing up of this study, which was originally my PhD thesis. First of all, my great thanks to Scott Lamb, sine qua non, who on his own time and with negligible remuneration, custom developed the tag counting and concordancing software (*LinguaStat*©) that allowed me to search and manipulate the corpus data. Without this application, it is certain that the analysis phase of the project would have lasted many years instead of only a few months. For their comments on drafts and helpful advice over the years, I give sincere appreciation to my PhD supervisors, Professor Jim Miller in the Department of Theoretical and Applied Linguistics (TAAL) and Professor William Gillies in Celtic Studies. Dr Wilson McLeod read a version of the grammar and offered many useful comments. Gordon Wells did the same, along with early versions of chapters 2–7. Daibhidh Grannd also offered errata and suggestions that have been incorporated into the present version of the grammar. Professor Cathair Ó Dochartaigh, as I embarked on the study, offered advice on corpus research and painted a realistic picture of the PhD process. The MacDonald and MacIsaac families, of Croismoraig, Cnoc na Lùib and Cladach Chircebst, North Uist, provided great help with learning Gaelic, acted as informants, and offered unceasing encouragement. Donald Morrison, Angus MacDonald, Iseabail MacLennan, and Mary Anne Kennedy at *Radio nan Gàidheal* were eminently forthcoming with transcripts, tapes and information about policy and practice in Gaelic radio. Rhona Talbot, at the School of Scottish Studies, provided the traditional narrative recordings and Dr John Shaw provided various transcriptions. Mairi Kidd sent library materials from Edinburgh to my Uist outpost. My thanks to Kirsty MacDonald for her help on transcriptions, and her patience and encouragement. Finally, my thanks to my daughter Ceit, who helped to quicken the mind and expedite the writing process over the last months before her impending arrival in 2002.

Map 1 is reproduced with the kind permission of Charles Withers. Map 2 was drawn by Colin Young.

William Lamb

University of the Highlands and Islands

Lewis Castle College

Benbecula

March 2008

Author's Note

This book is based on my University of Edinburgh PhD thesis, submitted in 2002. In that same year, I also published a version of Appendix 1 as the volume on Scottish Gaelic in the *Lincom Europa Languages of the World/Materials* series (Lamb 2002). When it emerged that the thesis had been accepted for publication by Cló Ollscoil na Banríona, John Kirk and I recognised that the framework of reference for many of the concepts and examples in the main body of the book necessitated the inclusion of Appendix 1 in the present volume and, moreover, that with the experience of use over the intervening years, numerous corrections and revisions to the grammar had become desirable. Appendix 1 may thus be regarded as a second edition of Lamb 2002.

Lamb (2005), an article which provided a blueprint for a large corpus of spoken Scottish Gaelic, drew on examples and some of the content of pp. 97–100 (§5.4.5: Juxtapositions and Subordination). Lamb (2006), which is a distillation of the results relating to noun phrase complexity and case marking incorporates, *verbatim*, paragraphs 1–4 from the current introduction (§1), paragraphs from the methodology chapter (mainly §4.1), material from §6.2.3, and a large part of Chapter 7. Otherwise, the rest of the present work has never been published before.

Transcription Practice

In this study, the spoken data have been transcribed primarily in accordance with the Gaelic Orthographic Conventions (GOC) of 1981. In some cases, the transcriptions have been modified to accord with the Gaelic Orthographic Conventions of 2005. In other cases, standard Gaelic has been deviated from to show certain dialect features of North Uist or to reflect the Uist speaking voice. One such feature is the shortening of the vowel in unstressed positions in *direach* ('just', 'exact(ly)') and *deanamh* ('doing', 'making'), reflected by the absence of a grave accent. Such departures from standard Gaelic orthography are coherent and logical. Where mis-spellings occur in written or printed texts, [sic] has been added editorially.

The presentation of examples

In the main text, there are many numbered examples comprising the Gaelic text and an English translation. The Gaelic is italicised; the English is in roman; any highlighted feature and its translation are in bold.

A great many examples also have a part-of-speech gloss. All three lines – the Gaelic text, the POS gloss, and the English translation – are in roman. The particular features in Gaelic under discussion are in bold.

In Tables 107–118, both the Gaelic text and the English translation are in roman. The particular features under discussion are in bold in both the Gaelic text and the English translation, and other special features already in bold are italicised.

Preface

Will Lamb found his way to my office one day in 1996. He had come to the Department of Celtic at the University of Edinburgh to take the MSc in Celtic. At the time, I was attached to the Human Communication Research Centre (HCRC), working on the syntax and discourse organisation of spontaneous spoken language. Will had become interested in the analysis of spoken Gaelic and wanted to choose a topic for his MSc dissertation that he could carry over to his proposed PhD research.

The MSc was successfully completed and Will began his PhD research. The original plan was to analyse the language of Gaelic news broadcasts on BBC Scotland (*Radio nan Gàidheal*), not just describing the language of contemporary broadcasts but also investigating who had made decisions about what Gaelic constructions and vocabulary could be used. Since news broadcasts in Gaelic were still relatively new and since the individuals who had taken part in their development were still alive, we hoped that Will's PhD thesis would offer an account of how a special Gaelic genre or text-type had been developed, elaborated and codified.

Will quickly realised that the enterprise required an account of spoken Gaelic but that no accounts had been written. It did not take much discussion before we decided that he should change the topic of his PhD thesis to the syntax of spontaneous spoken Gaelic. Through his former wife Kirsty, Will had access to native speakers of Gaelic who used the language in most aspects of their daily lives. He was able to record conversation at table, in the kitchen and at work. Not so much 'friends of friends', as in Lesley Milroy's Belfast fieldwork, but 'relatives of relatives' and 'friends of the family'.

Persuading people to converse while a cassette recorder is running is not always easy, and ensuring that the recording is of good quality requires skill and experience. But the hard slog begins only after the recordings have been made. Transcribing is exhausting and mind-numbing. Analysing the transcription brings its own difficulties: how many clauses are there in this chunk of speech? is this small phrase inside a clause or independent? what construction is this anyway? These are problems that plague native speakers of English working on English, but the difficulties are much greater for analysts who are not working on their native language and who have no extant grammars of the spoken language as landmarks. Will stuck to his task, in spite of the distractions of teaching, and produced an excellent PhD thesis. He has pioneered research on the syntax of spontaneous spoken Gaelic and, in doing so, contributed to the growing body of general work on the grammar and discourse organisation of unplanned speech.

Will's book has much to offer to all linguists, Celtic specialists or not, concerned with the grammar of spoken language. Given Will's intellectual, spiritual and practical commitment to the Scots Gaelic language and culture, his book will give especial pleasure to scholars of Celtic and of Scots Gaelic in particular.

Professor Jim Miller
University of Auckland

1 Introduction

1.1 Background and rationale

In many ways, a language's response to the contextual changes experienced by its speakers is a good gauge of its overall health. By surveying a large dictionary of a major world language, such as French or English, one can appreciate the wide range of words and phrases that have arisen in response to shifts in culture, technology and social mores. The more entrenched that a context becomes in a society, the more likely it is that it will be associated with a rich mode of linguistic expression. When new inventions, practices or other culturally important phenomena emerge, speakers quickly develop ways to communicate about them. Similarly, many of the obsolete words and expressions of a given language have suffered their fate as a result of their associated contexts ceasing to exist.

During the past several decades, a growing body of research (e.g. Ochs 1979; Chafe 1982; Biber 1988; Miller and Weinert 1998) has set out to delineate and contrast the characteristics of linguistic varieties associated with different parameters of context. Generally, these studies have compared two or more varieties of a language (e.g. formal prose and conversation) with regard to a number of linguistic features such as word length, use of subordination, contractions, construction-type and syntactic complexity. It has been found that linguistic responses to situation extend beyond the lexicon, with the most robust markers of variation in fact obtaining at the level of morphosyntax. One recurring theme in these studies is the major difference in the linguistic make-up of unplanned, informal speech *versus* other language varieties, especially planned, written prose.

It is no coincidence that the majority of these studies have been carried out on English, which is in an undeniably healthy linguistic state. There are few contexts or communicative purposes extant for which English does not have a well-developed vocabulary and idiom. As numerous studies in addition to the aforementioned have demonstrated, there are specific and detectable ways that language varieties associated with context – or *registers* – differ in the English language. However, the situation is not so well-described for most of the world's minority languages, nor is it universally accepted that registers exist in all of these.

Some researchers, such as Dressler (1988), have asserted that one of the qualities of a dying language is monostylism – the lack of register variation. It is reasoned that as languages 'die' they tend to be used for fewer and fewer contexts until they retreat to either a purely domestic or ritual domain. The process of language extinction is one of mass obsolescence; instead of words and idiomatic expressions falling into disuse gradually, in tandem with the situations that gave rise to them, entire contexts become linked to a higher-status language in lieu of a lower-status one. One might claim that Gaelic, for many of its speakers at least, is now in such a state. Because virtually all speakers of Gaelic are bilingual and surrounded by English on every front, there are almost no situations or functions for which the language is necessary and increasingly fewer for which it is reliably chosen as a means of communication.

It was in this context that I set out to examine the extent of grammatical variation in Gaelic. Although Gaelic seemed to me to still possess a relatively high level of variation, at least in certain domains, there were no data available for determining how much and what types of variation existed. I had conducted an earlier study into the linguistic variation of radio news-speak (Lamb 1999), which traced its historical development and attempted to associate changes in its context with ones in its linguistic form. However, despite the emergence of a number of interesting findings, I was limited

in the extent to which I could comment on what lexico-grammatical features distinguished news-speak from, say, conversation, fiction or formal prose. Additionally, as the study was primarily qualitative, I could not quantify the changes that had occurred in certain features over the years,¹ nor whether these changes were, in fact, statistically significant. The shortcomings of the study highlighted the need to investigate register variation in Gaelic at large, using a quantitative, statistical and computerised corpus model.

By using a quantitative model, it is possible to state the differences between groups of data in terms of relative proportions. Using the news-speak study as an example, this would have enabled a comparison in the occurrence level of different linguistic features at different stages of its development. By utilising inferential statistics, such as the Analysis of Variance (ANOVA), it becomes possible to state whether or not any apparent differences are significant at a reasonable certainty level, or on the other hand more likely to be due to chance. While the above study was corpus based as such (it used a finite body of purposely-collected text as its data source), by compiling *computerised* corpora, researchers have the ability to search, tag, and manipulate their data to an extent far beyond that feasible with pieces of paper. After the completion of the news-speak study, it was clear that the next step in the investigation of Scottish Gaelic register variation would be to assume a methodology with these three qualities.

Biber (1988) is an example of such an investigation on English. As his data, he used two large computerised corpora of spoken and written English divided into 23 registers.² By electronically 'tagging' a large number of register-sensitive linguistic features in these corpora, Biber was able to achieve a detailed, quantitative comparison of the genres involved. Because the corpora that he used were representative of the wide range of contextual and functional variation in the English language, he also provided a benchmark of its synchronic linguistic variation. His most important findings were: that register variation was multi-dimensional – there were no universal distinctions to be found linked to any one contextual parameter such as mode (speech *vs.* writing) – and that linguistic features should not be viewed in isolation, but as bundling together into several co-varying groups with shared functions or conventions.

Since this study, other research has indicated that typical speech and writing are more divergent from one another than Biber (1988) would suggest. For example, all of the variation linked to mode in Greenbaum, Nelson and Weitzman's (1996) study of subordinate clauses, was attributed to one register: conversation. Chafe (1992, quoted in Chapter 2) emphasises that informal conversation has the potential to tell us more than any other register about human language and psychology. Miller and Weinert (1998), a wide-ranging survey of spontaneous speech in numerous languages, detected a number of literate constructions that did not occur in their corpus of spontaneous speech, and colloquial constructions that were never found in their corpus of writing. They suggested that Biber failed to uncover large differences between the modes in his study because, as he himself acknowledges (1986), his spontaneous spoken samples had been recorded from academics: the segment of the population whose speech is most likely to be influenced by heavy exposure to and faculty with formal, written prose. Clearly, any study of a language's register variation must include a reliable and valid sample of on-line, informal speech. As will be discussed later, this is precisely the variety of Scottish Gaelic which researchers have the least access to and, consequently, know the least about.

¹ Such as passive forms. Over the course of the past fifty years, there has been an apparent move away from synthetic passives, such as *rinneadh e* 'it was done', to analytic ones, such as *chaidh a dhèanamh* (same gloss).

² These were the Lancaster-Oslo-Bergen Corpus of British English (the LOB corpus) and the London-Lund Corpus of Spoken English

In developing their approaches and interpreting their results, the aforementioned researchers had at their disposal several decades of work on English register variation and descriptive, functional syntax. In contrast, there has been scant research devoted to Gaelic register variation on either side of the Irish Sea. Of the studies that have appeared³ few have offered empirically garnered evidence, and none have drawn from a range of different registers. Their conclusions have been either of a general nature and readily extrapolated from other minority language situations, or based purely upon informal observation. To summarise these: *formal registers* evince less language contact phenomena, have a more conservative morphosyntax, and take their cues from the currency of institutionalised higher discourses such as religious language and oral literature; *informal registers*, on the other hand, in addition to displaying qualities converse to those above, also have a higher frequency of 'colourful language' (e.g. insults and asseverations). These are general characteristics and far from specifically pertinent to Scottish or Irish Gaelic. There are no discussions of base morphosyntactic features – such as the TAM system (tense-aspect-mode), focus constructions, or case marking – and no reference to categories of lexis. Furthermore, formality is the sole axis of variation discussed in these studies.⁴ There is no coverage of the differences of speaking and writing and no finer-grained examination of the role of situational context.

Similar to the situation regarding register research on Scottish and Irish Gaelic, very little work has appeared on syntax as opposed to that on phonetics and phonology (Macaulay 1979; Ramchand 1995). In Scottish Gaelic, there has been especially little interest shown in the morphosyntax of spontaneous speech.⁵ The effects of this dearth can be perceived in available ScG grammars, most of which are outmoded and make only passing reference to the spoken language, if at all. Perhaps associated with this, few pedagogical materials provide any information on the ways in which Scottish Gaelic speech diverges from writing. This prescriptivism has resulted in a large number of Gaelic learners who have been given the tools for writing formal Gaelic, but not for participating in colloquial speech situations.⁶

The formal syntactic work that has appeared on Scottish Gaelic has been almost exclusively generative in nature, with its application to questions of function and register necessarily limited. Within generative frameworks, Gaelic has more often been employed as a theoretical proving ground instead of a system still requiring description for its own end. Generative work tends to base analyses on highly planned, decontextualised structures more characteristic of writing than speech, and, often, 'magnasyntax' as opposed to what one would expect to find within any one individual's linguistic competence (Biber 1988: 7; Miller and Weinert 1998: 377).⁷ Grammatical treatments in these studies have been insensitive to questions of mode and register, providing categorical rules for morphosyntactic phenomena that become flimsy upon a proper consideration of these issues. This is no surprise, as conventional Chomskyan-type approaches to grammar do not admit different grammatical systems in the same language.

³ *Scottish Gaelic*: Dorian 1994, 1994b; Macaulay 1982a, 1992; Meek 1990; *Irish*: Ní Laoire 1988, 1993; peripherally Ó Siadhail 1989

⁴ Ní Laoire (1988: 90) defines register as a single, diglossic dimension based upon formality, as in other research of a more traditional orientation.

⁵ Ní Laoire (1993) describes a similar situation for Irish.

⁶ Macaulay (1979: 28) says: '... pedagogic grammars of Gaelic, and indeed Gaelic pedagogic materials in general, have concentrated on the written language ...'. However, in the same article he claims to make use of a corpus of spoken texts in his examination of prosody, but later admits that 'I have used a very restricted range of constructed sentences to exhibit the tone distinctions, rather than the *very heterogeneous material in my actual data texts*, which the reader might have found *intolerably confusing*' (p. 37) [italics added for emphasis].

⁷ There is also a confound in these studies arising from researchers not being fluent speakers of the language and having to rely on translation prompts for their informants.

To sum up, in approaching a study of Scottish Gaelic register variation, there are a number of obstacles to contend with:

- The research conducted thus far on register variation in Goidelic languages has been largely anecdotal and confined to a single axis of variation (formality).
- There are no studies of Gaelic style, genre or register utilising base morphosyntactic features.
- There is virtually no morphosyntactic (or discourse analysis⁸) work devoted to spontaneous spoken Gaelic.
- Of the little work that has been done on Scottish Gaelic syntax, most is generative in orientation.
- There is a concomitant lack of resources such as tapes of free conversations or grammatical descriptions based upon the spoken language.

Despite these issues, researchers such as Besnier (1988) and Jang (1998) – working on Nukulaelae Tuvaluan and Taiwanese respectively – have managed to conduct corpus-based register studies on languages presenting similar difficulties. And, there is a substantial body of work on register variation and the spoken syntax of the larger European languages (English, French, and Russian in particular) from which one may, assuming a degree of universality (see Biber 1995), begin to extrapolate. Therefore, there are precedents for successful register studies on languages with as few resources as Gaelic, and established models to follow based upon more researched languages.

Although empirical work on Gaelic register variation has not yet appeared, it has been called for by several authors:

An approach which would complement the one adopted here [a transformationally-driven examination of Gaelic clause structure] would be to study the relative frequency of occurrence of the various construction types discussed. To be useful, such a study would of course need to differentiate between texts from ... different stylistic varieties, most obviously between the spoken and written language. (Cram 1984: 47)

... [I]t is important to note that further categories of linguistic features are available [to be researched] in Irish [or ScG, *mutandis mutatis*] for the purpose of stylistic variation. Those which involve manipulation of syntactic structure and word order may be mentioned in particular. (Ní Laoire 1988: 289)

Relatively little has been published on ... the Gaelic language in its contemporary social context, and for that reason the article [on register variation] will be more tentative and more programmatic than the writer would wish. (Macaulay 1982a: 27)

A good deal of research needs to be undertaken in the area of participant relationships and its language reflection in general. (Macaulay 1982a: 40)

The research programme outlined by these statements would be quantitative: “study the relative frequency of occurrences” (Cram 1984); involve various construction types, particularly those concerning syntactic and word-order variation (focus, etc.); incorporate texts from a number of different registers; examine language in a contemporary social context; and investigate the role of participant relationships. Together, these statements constitute a nearly perfect abstract of Biber (1988) and also describe the orientation of the current investigation.

⁸ Dorian (1999) is, to the author's knowledge, the only study on Scottish Gaelic examining linguistic phenomena above the level of the clause.

1.2 Research Questions

Since there had been little empirical work on Scottish Gaelic register variation, the present study was an attempt to move towards a fuller understanding of how linguistic form and contextual variation interact in the language. It particularly attempted to highlight the characteristics of spontaneous, spoken forms, which have received little attention in the syntactic literature.

A computerised corpus was compiled of 81,677 words total, and distributed over four written and four spoken registers. These are as follows:

Spoken Sub-corpus

1. Conversation
2. Radio Interview
3. Sports Broadcast
4. Traditional Narrative

Written Sub-corpus

5. Academic Prose
6. Fiction
7. Popular Writing
8. Radio News Scripts

The texts were hand-tagged for around 100 discrete variables to reveal fluctuations in morphosyntax, information structure, and lexical class. By providing a thorough description of the contextual aspects of each register and investigating the functions of the linguistic features, it was possible to analyse the conventional associations holding between them.

The findings are important not only for our knowledge of Gaelic grammar and discourse, but can be placed in direct comparison to the findings of research on register variation and spontaneous speech generally. However, as it was the first empirical study of contextually based linguistic variation in a Goidelic language and the first to examine in-depth the grammatical characteristics of spoken Gaelic, it has particular relevance for Gaelic studies. It also put Dressler's prediction to test for this particular endangered language.

Such a study must rely upon a functionally aware grammatical description. As mentioned previously, available grammars of the language are almost devoid of any description of spoken Gaelic. Furthermore, some early publications are parasitic upon Latinate paradigms, especially in their description of case and verb inflection. It seemed necessary to me to write a new descriptive grammar of the language to inform the study, particularly during the tagging process. This effort soon grew in size to the extent that it became somewhat independent. I published it with Lincom Europa (Lamb 2002), but since it has been available up-to-now only at an exorbitant expense, and since I am under no binding contract with the company, I offer it here, in its original context.

The specific research questions examined in this study are as follows:

1. What linguistic features are significant markers of Scottish Gaelic register variation?
2. What contextual features help to condition linguistic form in Scottish Gaelic?
3. What associations between context and linguistic form are present in the corpus?
4. What are the main grammatical characteristics of spontaneous spoken Gaelic?
5. How do the findings relate to relevant research on other languages?

1.3 Assumptions and Limitations

As in Biber (1988 *inter alia*), Ferguson (1983: 1994) and generally all studies of register variation, it is assumed that communicative contexts that recur habitually in society are associated with conventionalised linguistic features that diverge from the linguistic features of other communicative contexts. Also, it is assumed that linguistic features which co-vary across a particular set of contextual parameters share at least one common communicative function. Following from this assumption, it is justifiable to interpret groups of co-varying features via the communicative functions associated with their underlying contexts.

In terms of limitations, with only eight registers totalling 81,677 words implicated in the current study, it is not possible to achieve a complete picture of the Gaelic language. However, the registers have been carefully chosen to be characteristic of the different kinds of discourse found in the language, as well as to maximise the potential for a wide range of linguistic variation. It should be said that dialectal variation has not been controlled for, beyond using mainly Uist speakers in the conversational texts, and that this is a potential confound. Macaulay (1992: 153) states that the Gaelic dialects are syntactically “remarkably homogenous”, yet as a study of variation in Gaelic syntax – dialectal, diachronic, or otherwise – has still to be done, there are no grounds for assuming that there are not differences. Also, authorship and participant age were not controlled for beyond only using texts that were written or collected post-1965, and speakers who were old enough to be considered fluent in the language. Gender effects have also not been controlled for. However, in light of the time constraints imposed upon the study, and the difficulty in obtaining spoken Gaelic texts, I hope that the reader will excuse the above limitations.

1.4 Definitions

(The) corpus: refers to the computerised database consisting of 81,677 Scottish Gaelic words from eight registers compiled for this study

Features: these are the discrete linguistic variables (e.g. attributive nouns) that were counted and compared across the eight registers (listed in §4.2)

Mode: speech or writing

Register: “a general cover term for all language varieties associated with different situations and purposes” (Biber 1994: 32); also subsumes the terms *genre*, *style*, and *language/linguistic variety*

Tag Set: the collection of ‘codes’ representing each of the discrete linguistic variables used in the register comparisons (see Appendix 3 for a full list)

2 Spoken and Written Registers and Scottish Gaelic

2.1 Speech and Writing

Although the emergence of empirical research on the subject occurred properly only in the late 20th Century, the concept of a differentiation between speech and writing has a prolonged history. Academics at large have long viewed writing as the purer form, free from the disfluencies and fickleness that characterise common speech. This, what we could call the *aesthetic* orientation towards speech and writing, began to wane in linguistics in the 19th century and was largely supplanted by another view. As Biber (1988) remarks, with the initiation in the 19th century of phonetics as a separate discipline, spurred on by the work of Jacob Grimm, the German folklorist-linguist, and later, Henry Sweet and Daniel Jones, language researchers began to turn their attention more towards speech as an object of inquiry. The 'Brothers Grimm' had set out to collect the natural, colloquial speech of their informants. Previous collectors had been content with, or even idealised, literary versions of stories that had originally come from oral sources. Folklorists throughout the world took up the Grimms' example and left us with invaluable transcriptions, not only of folklore, but also of the lexis and grammar of many long-disappeared or irrevocably changed language groups. John Francis Campbell, or *Iain Òg Ìle* as he is known in Gaelic, was one of these, and he deserves recognition as the first spoken language researcher in Scottish Gaelic. In the introduction to his *Popular Tales of the West Highlands* (1994, first published 1861), he states:

I requested those who wrote for me to take down the words as they were spoken, and to write as they would speak themselves; and the Gaelic of the tales is the result of such a process ... The Gaelic then is *not* what is called "classical Gaelic". It is generally the Gaelic of the people – pure from the source. (p. 80)

With a growing interest in speech came the view in linguistics and anthropology, promulgated in particular by Bloomfield and Sapir, that spoken language should be regarded as primary and that written language was but an inferior artefact, particularly in terms of its status as a research focus. Bloomfield (1935) advances his position through certain facts about language: speech is a human inheritance while writing is a cultural one; to understand writing one must understand speech but not the converse; and in earlier times, literacy had been common only in a relatively small, privileged section of society. Sapir (1974: originally 1933) says the following about speech *versus* other kinds of human communication:

[T]he actual history of man and a wealth of anthropological evidence indicate with overwhelming certainty that phonetic language [i.e. speech] takes precedence over all other kinds of communicative symbolism, all of which are, by comparison, either substitutive, like writing, or excessively supplementary, like the gesture accompanying speech. (p. 46)

Later researchers have found other reasons to regard speech, and especially conversation, as fundamental. Halliday (1994: xxiii-xxiv) says that speech is unconscious and in flux in a way that writing is not; it maps itself onto an ever-changing environment, demanding a greater reservoir of semantic and thus grammatical strategies to cope. Miller and Weinert (1998: 29) discuss the fact that spontaneous conversation – along with narrative – is the language variety that we heard most as children and the one that we use the most as adults. Chafe concurs:

Speaking is natural to the human organism in ways that writing can never be. It is plausible to suppose that humans are “wired up” to speak and listen, that the evolution of speech was inextricably interwoven with the physical evolution of our species. The same cannot be said of writing ... If speaking has a priority in this sense, we can identify conversational language, as opposed to various manifestations of oral literature, as constituting the most basic kind of speaking ... Conversation can justifiably be taken as the use of language to which humans are best adapted and thus the one that can tell us most directly about inherent properties of language and the mind. (Chafe 1992: 88–9)

These two viewpoints, the *aesthetic* orientation and the *primacy* orientation towards speech and writing, continue to co-exist today. Outside of linguistics, it is clear that the aesthetic orientation still dominates; in business, school, and lay opinion; writing is the mode that holds official sway, and that encapsulates what is ‘good’ and ‘proper’ – our prescriptive views – about language.

Modern linguists continue to espouse the position that speech should be regarded as primary, yet there is a fracture between what many of them declare and what they actually do. This is particularly true in generative grammar, but is a general characteristic of all theoretical models (Miller and Weinert 1998: 4–5). With its emphasis on locating an underlying logical substratum to human language and testing the limits of this universal grammar, generative work often uses data that is more representative of the particular complexities of writing – refined and explicit – than speech, which we are supposed to be taking as prime. The hazard of this paradigm is that it is possible to overlook certain regular constructions in spoken language which simply do not appear in the written language when one is too ready to consign data to the realm of performance errors. Additionally, the use of intuitions as a data source may be a major confound: “[they present] only indirect evidence, as they reflect a variety of memory and processing constraints, as well as subjects’ efforts to conform – consciously or not – to prescriptive grammar” (Birner and Ward 1998: 27). The work of Biber, Chafe, Halliday, Miller and Weinert and others analysing natural spoken language has demonstrated that there are certain constructions that regularly appear in one mode but not the other. However, the aesthetic orientation towards speech and writing is still influential in modern linguistics.

2.1.1 *Linguistic Differences between Writing and Speech: Empirical Studies*

For the past 35 years or so, there has been a small but steady stream of research, often by linguists with roots in functional grammar or linguistic typology, which has attempted to delineate the differences of written and spoken language. As will be shown, researchers have often arrived at different conclusions depending on their data and method of analysis. This section will summarise earlier investigations on the subject, examine contemporary work, and conclude with a consideration of register, moving away from the often problematic emphasis on mode.

Blankenship (1962) conducted one of the first empirical studies on linguistic differences due to mode, comparing the spoken lectures and published monographs of four public individuals. She found that idiolect accounted for more of the group variation than did mode; that there was more similarity than dissimilarity between the spoken and written texts. (However, the effects of planning time, as investigated by Ochs (1979), may be implicated in this; unlike spontaneous conversation, lectures are at least semi-planned.) Her oral texts exhibited more imperfect aspect, less present perfect aspect, and less passive constructions. There was no use of inferential statistics and, consequently, it is not clear whether the differences obtained were due to chance alone. Blankenship (1974) was more complex, comparing six individuals’ spoken and written texts across six different sub-modes (registers). She examined features such as sentence length, word length, the

type-token ratio⁹, and the frequencies of different parts of speech, as well as employing inferential statistical tests. As in her previous study, she found idiolect to be a more robust determiner of linguistic form than mode, but she did obtain significant differences between the modes and registers on a number of features.

O'Donnell (1974) used a sample of spoken and written discourse produced by a single public figure. He divided the sample into t-units (an independent clause along with any related dependent clauses) and counted a number of features including different kinds of clauses, nonfinite verbals, passive constructions, auxiliary verbs, and attributive adjectives. Frequencies were subjected to Chi Square tests and his results indicated that writing was more syntactically dense (indicated by having shorter t-units) and had more dependent clauses, gerunds, participles, passives, attributive adjectives, and modal/perfective auxiliaries. The spoken sample had a greater instance of noun clauses, infinitive constructions, and progressive aspect.

Poole and Field (1976) used a structured interview and timed written exercise from 80 undergraduate students discussing life experiences. The samples were then compared on a number of linguistic features such as prevalence of subordination, number of different parts of speech, mean sentence length, and passive verbs. Inferential statistics were used and significant differences were found between the modes on a number of features: there was more adverbial elaboration in speech but more adjectival elaboration in writing; there were more passives in writing; there were also more 'preverbal' constituents in writing (e.g. preposed subordinate clauses). In contrast to many other researchers (save Halliday), they concluded that oral language was more structurally complex than written language. However, the results may be queried on the basis of three qualities which might have skewed the study. First, the elicitation instrument that was used, a 'structured interview schedule', would have conditioned a more formal setting than is typical of most spoken language production. Second, they employed a group of participants – tertiary students – who, through their intense consumption of and proven ability with formal, expository writing, would have been more likely than the population at large to use literate features in their speech. Third, as the written exercise was timed, it introduced production constraints more typical of oral than literate language production. Therefore, the context underlying the spoken data in the study was more typically *literate* while the context for the written data was more typically *oral*.

Ochs (1979) was interested in ways in which adult spontaneous speech often resembles early child discourse. She examined a wide range of language data involving both adults and children but, in this article, focused on informal conversations between adults and personal narratives from single speakers in two conditions: spontaneous-spoken and prepared-written. Spontaneous spoken discourse was characterised by the features in the following list, many of which are similar to the properties of early child speech. Planned discourse was characterised by the converse of the below and generally exhibits the more complex morphosyntax that humans acquire later in life:

1. Speakers express propositions through involving the immediate context (e.g. by pointing, lifting objects, or gazing).
2. Semantic relations between referents and following predications are not explicit (e.g. as in left-detachment constructions).
3. Conjunction of propositions is found without explicit marking (asyndesis).
4. Demonstratives are used in lieu of definite articles or relative clauses (e.g. 'this thing ...').
5. Low frequency of passives.
6. Use of 'historical' present tense.
7. False starts and word replacement and repair mechanisms.
8. Phonological/ lexical repetition and elaboration.

⁹ The proportion of different words (types) to the overall number of words (tokens) in a text; a measure of lexical variety.

Using a similar data set, Tannen (1982b) examines the spoken and written versions of narratives produced by two individuals. Her findings are in line with many of Ochs':

1. Written versions tend to stay in past tense while spoken versions switch to the historical present.
2. Spoken versions tend to use different kinds of determiners.
3. Written narrative is more concise.
4. While written narratives use explicit marking of related propositions, spoken ones tend to rely on juxtaposition.
5. Spoken versions tend to convey meaning through syntax and the specific 'rhythm' that it entails – in writing, meaning is lexicalised.
6. Cohesion is formed in writing by syndesis, lexicalisation, and 'focus' constructions while it is formed in speech through paralinguistic strategies (pointing, laughing, yawning, etc.).

However, Tannen takes the position that register, rather than planning time (as construed by Ochs) or mode, is primarily responsible for these kinds of differences. She also points out (later argued forcefully by Akinassio 1985) that speech and writing can resemble each other in various ways. For example, short stories, while presenting tight, integrated prose, can simultaneously enlist involvement strategies (such as direct quotes, and sound/word repetition) in an attempt to bring the reader closer to the text. In fact, some imaginative prose goes so far as to mimic the typical disfluencies of spontaneous speech.

Chafe's (1982) influential paper is the most detailed empirical investigation of its time on the contribution of mode to linguistic divergence. He summarises the main findings from early research on the overall differences between speech and writing as follows:

Compared to Writing, Speech has:	Author
Longer texts	Drieman 1962
Shorter words	
Less attributive adjectives	Devito 1966, 1967
Less vocabulary range	
More words referring to the speaker	
More quantifiers and hedges	
Less "abstractness"	
Less subordinate clauses	Harrell 1957; O' Donnell 1974, Kroll 1977

Building on the findings of these studies, he investigates a corpus of the language of academicians, comprising 9,911 words of informal speech and 12,368 words of formal writing. Chafe was careful to emphasise that he was sampling extremes of mode, and that some forms of writing were closer to more typical speech and vice versa. He found that he could place the differences detected along two dimensions associated with different levels of orality and literacy: fragmentation *vs.* integration and involvement *vs.* detachment.

Fragmentation, an aspect of spoken language, was distinguished by features such as asyndesis and the presence of initial co-ordinators (*and, but, or*, etc.) while *integration* – the tight packing of information in careful writing – was indicated by nominalisations, genitive subjects and objects, participles, attributive adjectives, and certain kinds of so-called subordinate structures. *Involvement* referred to the manifestation of speakers' association with their audiences, appearing for instance in higher frequencies of first person references, references to speakers' own mental processes, discourse particles, emphatic particles and hedges. *Detachment*, on the other hand, was characterised as linguistic distancing from "specific concrete states and events" (p. 45) and was characterised by passives and nominalisations. Chafe also made the important observation, using ritual texts from the Seneca language, that these dimensions seem to

characterise the difference between oral literature and colloquial speech as well. Later, Akinnaso (1985) explored this area in more depth, comparing and contrasting Yoruba divinatory speech, formal writing, and colloquial discourse.

Chafe and Danielewicz (1987) are perhaps alone in reporting a "relatively pure difference" (p. 91) between speech and writing that is not affected by register. They found that speech was characterised by hedges, contractions, disfluencies, reduced lexical range, a higher frequency of coordination of intra-clause constituents, and one-word sentences. They also found that spoken language more frequently evinced simple strings of coordinated clauses, avoiding complex intraclausal relations.

In a series of papers, researchers Sidney Greenbaum, Gerald Nelson and Michael Weitzman have investigated the distribution in the *Leverhume Corpus*¹⁰ of a number of linguistic features suspected of being sensitive to modal variation: clausal relationships (Greenbaum and Nelson 1995a); nuclear and peripheral clauses (Greenbaum and Nelson 1995b); positions of adverbial clauses (Greenbaum and Nelson 1996); complement clauses (Greenbaum, Nelson and Weitzman 1996); cleft constructions (Nelson 1997); and elliptical clauses (Greenbaum and Nelson 1999). Greenbaum and Nelson (1995a) found that, although instances of subordinate and coordinate clause combinations occurred more frequently in written samples, there were wide fluctuations of occurrence levels within each mode. For instance, while conversation was found to have the lowest proportion of coordinate and subordinate clauses, spoken monologue had the highest. They concluded that mode, in and of itself, was not a reliable predictor of variation in clause combinations. Similarly, in a replication of an earlier study on subordination (Thompson 1984), Greenbaum and Nelson (1995b) found that both mode and formality were implicated in the distribution of peripheral modifying clauses. While more typically written, formal registers had higher levels of peripheral clauses overall, informal spoken registers (i.e. conversation and broadcast discussion) evinced the highest proportion of finite adverbial clauses. A later study (Greenbaum, Nelson and Weitzman 1996) analysed the distribution of all types of subordinate clause in the *Leverhume Corpus*. It revealed that the conversational data accounted for *all* of the variation attributed to mode. This result adds weight to Miller and Weinert's (1998) conclusion that spontaneous speech, and in particular conversation, is radically different from other forms of language. However, the same study, when examining complement clauses alone, found robust differences between the modes: in speech most complement clauses were finite while in writing, most were non-finite. Nelson's (1997) study on clefts failed to detect large differences between speech and writing and he concluded that they ought to be investigated at a more specific discourse level. Greenbaum and Nelson (1999) revealed differences between speech and writing in various types of clausal ellipsis, but also found that certain registers were aberrant and exacted a heavy influence on the results: "... our results for unscripted monologues, and so for speech generally, have been heavily influenced by the sports commentaries ... [which] probably deserve a separate study" (p. 124).

In summary, most empirical studies have located few strict differences between the modes. As mentioned previously, one of the problems in investigating 'mode' involves which *registers* are being taken as representative of speech and writing. If one researcher, for example, includes spontaneous commentary in an investigation of spoken and written clausal ellipsis while another does not, their results will be unequal. In light of the tendency for certain registers to be considerably divergent from other ones within a given mode, it seems preferable to appeal to a finer model of linguistic and contextual variation. One solution to this issue is presented in §2.3. It is also possible to carry out a close examination of the discourse properties of a single language variety. This kind of approach has been taken by researchers such as Miller and Weinert (1998) in their analysis of spontaneous spoken language (see §2.1.3 below). Yet, another problem faces researchers in this field: to what extent are the linguistic constructs that they employ actually valid

¹⁰ A sub-corpus using written and spoken texts taken from the British section of the International Corpus of English (ICE).

and consistently defined? The next section examines the implications of this issue for two areas that have received more attention in the literature than any others: subordinate clauses and the notion of linguistic 'complexity'.

2.1.2 *Complexity and Subordination in Speech and Writing*

In general, studies have concluded that writing is more complex and more dense than speech although some researchers, such as Halliday (see below), have emphasised the different *kinds* of complexity inherent to speech and writing. Chafe's dimensions of integration and fragmentation characterise one aspect of complexity, attributed to the low time constraints generally found in writing and the potential, therefore, for greater cohesion: "the packing of more information into ... [syntactic units] than the rapid pace of spoken language would normally allow" (p. 39). Highlighting another contrast between the modes, Tannen (1982b) points out that speakers can achieve cohesion through non-verbal strategies and paralinguistics while writers must resort to grammaticalised connectives and complex syntax.

Most of the research on oral-literate differences has focused to some extent on the distribution and types of subordinate clauses in spoken and written language. Biber (1988) summarises the often confusing discrepancies in the literature associated with this area. Most researchers have found writing to contain more subordination (Kroll 1977; Chafe 1982; Brown and Yule 1983; Thomson 1984). In a like manner to Tannen (1982b), Mithun (1985) speculates that writing might place a greater need on subordination due to the lack of resources available in speech – such as intonation – for the bonding, backgrounding, and foregrounding of clauses. Ochs (1979) explains what she finds to be a greater frequency of coordinate and subordinate clause complexes in writing as a consequence of planning time.

On the other hand, researchers such as Poole and Field (1976) and, in particular, Halliday (1987; 1994) have claimed opposite findings, that speech is more syntactically rich than writing. Halliday maintains that writing and speech have different kinds of complexity, speech being more intricate in the way it associates clauses with one another and writing being more rich lexically. For instance, what writing accomplishes through a dense lexicon, speech accomplishes through coordination, subordination, and the simple juxtaposition of related clauses. This is similar to Chafe and Danielewicz's (1987: 92) comment that speaking achieves its richness through the steady modification of a limited range of choices while writing broadens this range.

Beaman (1984) sees the disparities in the literature concerning subordination as being due to a conflation of mode and register, there defined as formal *vs.* colloquial discourse. She adheres to one register – narrative – and uses spoken and written forms of it to explore inter-modal differences in the types and frequencies of subordinate clauses. She finds a greater number of finite subordinate structures (such as relative clauses) in speech while finding more non-finite subordination in writing (e.g. participles). In addition, she finds a higher number of coordinated sentences in writing (38% *versus* 25% in speech), but concurs with Halliday that subordination is, overall, more frequent in speech (18% of sentences *versus* 13% in writing). Greenbaum, Nelson, and Weitzman (1996) received similar results in their corpus: they obtained more finite complement clauses in speech and more non-finite complement clauses in writing. However, Greenbaum and Nelson (1995) criticise Beaman's research as confounded by not representing natural discourse (the texts were experimentally elicited) and using only female subjects.

Another possible source of confusion in the literature about subordination may lie in the concept itself. Thomson (1984) – along with Haiman (1985) and Mathieson (1988) – questions its descriptive validity and advocates for excluding the term from linguistic research. She considers its traditional bases, such as evidence of dependency, ability to prepose, reducing, and backgrounding as "at best *ex post facto* rationalisations of our own (Western educated) 'intuitions'" (1984: 86) making them necessarily tautological. In attempting to hone in on a better definition, she disqualifies clauses which have

constituent relations with nouns, verbs, or prepositions (i.e. relative and complement clauses: those that 'embed'). This leaves adverbial clauses, participles and non-restrictive relative clauses (those that 'combine') and she states that "[these are not] more or less determined by the grammar of complementation and the pragmatics of reference" but rather "have... to do with how one decides to convey and relate propositions" (p. 85).

Instead of abandoning subordination as a working term, Van Valin (1985) *et al.* (Foley and Van Valin 1986; Van Valin and LaPolla 1997) attempt to further refine the notion, reallocating and stratifying the range of syntactic relations previously subsumed by it. For instance, in Role and Reference Grammar (most fully developed in Van Valin and LaPolla 1997), restrictive relative clauses are viewed as NP subordination while adverbial clauses are an instance of subordination at the clausal level. (For more information on Role and Reference Grammar and clausal relations, please refer to Appendix 1: §3.3.)

So, the disparities between studies on complexity and subordination issues may fundamentally involve inconsistencies of definition. Some studies have had different understandings of mode and register while others have disagreed on the very basis of subordination and linguistic complexity. Perhaps the most reasonable view of complexity is that of Halliday, who highlights the different ways in which both typical speech and typical writing are complex. The controversy about subordination is still being played out and the current study, instead of attempting to resolve this particular theoretical debate, concentrates on the functions and distribution of individual constructions which have been classed as subordinate.

2.1.3 *Spontaneous Spoken Language (SSL)*

Both linguistic hinting and intuitive notions concerning the fundamental status of conversation in human language (see Chafe's quote in §2.1) have lead some researchers to further examine the specific features that differentiate SSL from other types of language. Using a wide range of data from Scottish English, Russian, and German including natural conversations and Map Task data, Miller and Weinert (1998) is the most extensive study on the properties of spontaneous speech. They concentrate on informal, colloquial speech and, notably, set out to gather data from non-academics. This is in contrast to other studies (e.g. Chafe 1982; Biber 1988; any research using the Brown corpus) that, as Miller and Weinert point out, might have detected greater differences between SSL and other kinds of language had their conversational data not come from one of the sectors of the population most accustomed to the consumption and production of formal, expository prose.

They claim, in tandem with past researchers such as Chafe and Tannen, that SSL has certain key fundamental traits (list from Miller and Weinert 1998: 22):

1. Spontaneous speech is produced in real time, impromptu, and with no opportunity for editing, whereas written language is generally produced with pauses for thought and with much editing.
2. Spontaneous speech is subject to the limitations of short-term memory in both speaker and hearer: it has been said ... that the short-term memory can hold 7 +/- 2 bits of information.
3. Spontaneous speech is typically produced by people talking face-to-face in a particular context.
4. Spontaneous speech, by definition, involves pitch, amplitude, rhythm, and voice quality
5. Spontaneous face-to-face speech is accompanied by gestures, eye-gaze, facial expressions, and body postures, all of which signal information.

These qualities are then linked to several linguistic features such as: the careful staging of information in small 'bite sized chunks' such as seen in cases of NP-Clause structures (or *left-detachment* – see Lambrecht 1994); lower incidence of grammatical subordination;

fragmented and unintegrated syntax (cf. Chafe 1982); importantly, that the sentence is not a useful unit of analysis for informal speech (cf. Halliday 1989); certain patterns of constituent structure and head-modifier placement do not agree with those corresponding structures recognised by syntactic theory; SSL has less lexical variation; and certain constructions that appear in SSL do not appear in written language and vice versa. Their study then progresses to examine in detail the syntax and morphology of SSL in terms of clause structure, the noun phrase, and focus constructions as well as the means and motivations of the discourse organisation found in their corpora.

At the very heart of the difference between writing and speech is the notion of the sentence. In English and most European languages, this is the most basic unit of writing: the words between the first capital letter after a full stop and the following full stop. Spoken language, on the other hand – especially informal spoken language – usually does not conform well to this division. Although some researchers have continued to use the sentence in their work on spoken language, they seldom advance explicit reasons for doing so.¹¹ Many researchers such as Nivet (1989), examining the sociolinguistic markers of French natural conversations, have discarded the traditional notion of the sentence as not suitable for oral description. Miller and Weinert assemble the following arguments against its use in spoken language analysis (1998: 33–71): parsing transcribed speech into sentences can be a largely arbitrary business with a significant amount of disagreement between transcribers; sentence length can be seen to vary both diachronically and inter-culturally; some samples of SSL even seem to primarily consist of phrases rather than clauses; and there is an incompatibility between many basic structures of speech and those structures of writing. Notably, Chafe and Danielewicz (1987) defend the applicability of the sentence to speech, appealing to the fact that speakers usually produce utterance-final intonation which they consider to be analogous to the use of the full stop in writing. However, Miller and Weinert counter that most of the ‘sentences’ from speech that Chafe and Danielewicz evince resemble paragraphs rather more than sentence units as we know them today. Miller and Weinert see sentences as “low level discourse units relating to the organization of clauses into larger texts” (p. 71). They conclude that we cannot argue for the presence of these structures in spoken language except when a spoken register is more akin to prototypical writing than speech, as in formal lectures.

If the sentence is banned from the analysis of spoken language, the issue remains as to what could be used in its stead. Ideally, when comparing speech and writing, one would employ a unit shared by both. Halliday (1989) maintains that the clause is the unit of analysis most basic to speech. Macaulay (1991), studying the spontaneous speech of sociolinguistic interviews, used the clause as the basic unit of his syntactic analysis. Greenbaum and Nelson (1995), when marking up their corpus of spoken and written English, also opted for the clause rather than the orthographical sentence as a unit of analysis. Their rationale was that sentences are often seen to be rhetorical units rather than syntactic ones. Like Halliday (1994) and Miller and Weinert (1998), they made reference to clause clusters (or complexes) to account for parataxis and hypotaxis. However, although agreement and government are sometimes seen to cross clause boundaries, the densest clusters of these phenomena are actually found within clauses (Miller 2000). Lyons (1999) concurs:

Despite what I said or implied in such works as Lyons (1968, 1977a), I now take the view that clauses, rather than sentences, are the basic units of syntax (and especially of government and dependency) and that (if one retains the

¹¹ It is sometimes claimed that straight lexical transcriptions of spoken language can resemble, in Halliday’s words (1994: xxiv) “a dog’s dinner” and be off-putting to readers (cf. footnote 6). However, analysis and publication are two very separate issues. Any time standard orthographic sensibilities enter into spoken transcription for the purpose of analysis, as in the British National Corpus, it stands to reason that a potentially confounded layer of interpretation has been laid on top of what is an essentially different type of data. It is as if apples have been magically transformed to pomegranates.

traditional concept of the sentence) sentences are best defined in terms of clauses, rather than conversely ... (p. 152)

In the current study, the clause will be defined as it is in Miller and Weinert (1998: 77): "a syntactic structure that relates to one state or event and has a nucleus and core (some elements in the core being optional or obligatory depending on the nuclear predicate) and an optional periphery".¹² In addition, with the exception of cosubordinate clauses, all clauses in the present research will be finite, defined by the presence of a tense-carrying verb.

To conclude this section, analysts have detected a number of ways in which spoken language – particularly natural, informal conversation – differs from written language. This account has been a general overview and it has not gone in depth into any particular linguistic features that characterise this difference *per se*. This was deemed more appropriate for chapters 4–7: wherein the composition of the tag set will be discussed as well as the motivations for including particular individual features, most of which have received prior attention in the literature.

The present study, however, is not limited to examining the linguistic corollaries of mode. One of the most important findings from Biber's work is that there are no pure differences between speaking and writing *per se*; in fact he found more variation *within* the modes than between them. The next section will examine research on the correlates of contextual and linguistic variation (mode being just one facet of this): the field of register analysis.

2.2 Register: Context-based Linguistic Variation

As seen above, the tangling of mode and formality has been one major confound in some of the work on speech and writing. Mode and formality, however, are only two of the possible axes underlying linguistic variation based upon context. Researchers have employed a myriad of terms to refer to this notion; genre, text type, (speech) style, rhetorical mode, and register are perhaps the most frequently encountered (see Biber 1994 for an overview of these terms in the literature). Some researchers such as Ferguson (1994) and Biber (1988) have found it useful to distinguish between them on the grounds of whether, for example, texts are grouped on the basis of external criteria such as message types (genres) or internal properties such as linguistic form (text type). Other researchers have used these terms synonymously, referring in an inclusive sense to language varieties correlating with specific uses or situations. Atkinson and Biber (1994) prefer to use the term *register* in this way while Miller and Weinert (1998) prefer *genre*. Unless there is a need to be more specific about the classifying criteria being used, 'register' will be used in the current study as aforementioned to refer to a language variety linked with some situational context, purpose, or use characteristics.

Register studies examine naturally occurring language in the form of texts. Intuitions will not suffice, nor word lists, nor sentences in isolation of their original texts (Biber and Finegan 1994). Register studies are, in essence, empirically motivated, fundamentally concerned with real language and the contexts within which different varieties of discourse occur. As they involve collecting natural discourse, register studies are also *perforce* corpus studies. McEnery and Wilson (1996: 24) describe a corpus as "a finite-sized body of machine-readable text, sampled in order to be maximally representative of the language variety under consideration". Although not all register studies have made use of computerised corpora, this is a growing tendency. However, there is certainly an onus upon all register researchers to collect texts that are maximally representative of the registers with which they intend to deal.

A basic assumption in register analyses is that "a communicative situation that recurs regularly in a society... will tend over time to develop identifying markers of language structure and language use, different from the language of other communication

¹² See §3.3 in Appendix 1 for an illustration of the parts of the Scottish Gaelic clause.

situations" (Ferguson 1994: 20). The first methodological step of a register study is to provide a linguistic description of these markers, which can occur at many levels of linguistic description, for instance in the lexicon, phonology, morphosyntax, and intonation. Biber (1994) recognises two types of linguistic characterisation in register analysis: register markers, being features which are only found in certain registers (e.g. the umpire 'count' – the number of balls and strikes – in baseball); the different frequencies of core linguistic features such as noun phrase density, clause types, discourse devices, tense usage, and modification types. Although certain registers will be easily recognised by the presence of particular words or phrases such as *one ball, one strike*, Biber tells us that lexical choice alone is not usually very helpful for placing registers.

Equally as important as the linguistic description is a full consideration of the contextual factors of each individual register. Once these two steps have been addressed, the analyst then attempts to specify the functional and conventional associations between linguistic form and situational context. This should be framed in terms of correlation rather than cause; situation should be seen to influence language and language should be seen as helping to give rise to the situation.

In the literature, there are generally two kinds of register studies: those that describe a single register – often its genesis and development – and those that focus on register variation, using a representative sample of the registers in a given language and showing how linguistic features and context co-vary in that language. Single register studies (see Atkinson and Biber for a literature review) – such as Ferguson (1983), Ghadessy (1988), Romaine (1994), and Lamb (1999) – have tended to describe the linguistic features, situational context, and functions detected in a limited language variety and are not studies of variation as such.¹³ In general, they have made little or no use of statistics and have focused on media language. On the other hand, by virtue of the fact that register variation studies are concerned with many different kinds of text and surveying a *language's* output rather than the constituents of a single contextual domain, they have tended to be statistical and data intensive. It is no accident that they have only come of age since the advent of quick-access computer corpora and high-speed tagging and searching programmes. However, the time required in assembling representative corpora, not to mention tagging it for linguistic features, has meant that there have been few investigations of this nature to date. Douglas Biber's work on English (1988) is the most well known study in this field. His methodology has also been applied to a small, but diverse, number of other languages including Nukulaelae Tuvaluan (Besnier 1988), Korean (Kim 1990), Somali (Biber 1995), and Taiwanese (Jang 1998).

2.3 Multi-dimensional (MD) Analysis

Biber (1988: 52–3) identified a number of methodological problems detracting from most previous studies on speech and writing differences. In general, they concern limitations in being able to account for the overarching differences of speech and writing:

1. Giving disproportionate weight to individual texts
2. Giving disproportionate weight to certain genres, with most studies only examining one or two at the most
3. Involving only a limited number of linguistic features
4. (in footnote) Involving only a select range of speakers/writers, often middle-class academics
5. Using only English data, but drawing conclusions about 'speech' and 'writing' in general

Jang (1998) adds to this list the lack of multivariate statistics, which give a fuller view of the variation when employing several variables. However, the majority of previous work does not even include inferential statistics; there are few exceptions in the literature that go beyond descriptive statistics such as mean and range. The limitation of these measures

¹³ Unless of course they happen to be diachronic studies.

is that the researcher cannot comment as to whether or not seemingly notable tendencies are merely due to chance. This fact begs the question whether a meta-analysis of previous studies would reveal that some accepted findings – and aberrations, such as the disagreement over levels of subordination – are actually statistically non-significant.

Using multivariate statistics including factor analysis, Biber's (1988) study on English attempts to move beyond the limitations of previous studies. He uses a representative corpus of spoken and written registers (N=23), from which he extracts a small number of fundamental dimensions of functional variation to account for the range of communicative situations and purposes available to a language. The dimensions have three crucial aspects (Biber 1994): none of them alone is able to account for all of a language's linguistic variation; they are scales of variation rather than binary oppositions; and they result from a quantitative analysis rather than a preordained functional one. Dimensions consist of bundles of linguistic features that co-vary in texts and are interpreted as sharing at least one common communicative function. The dimensions are used to characterise different registers with each variety of language scoring differently upon them, evincing different levels of each dimension's characteristics. Take Biber's (1988) Dimension 2 which is made up of positively loading features such as past tense, third person pronouns, and public verbs, and negatively loading ones such as present tense, attributive adjectives, and word length. Examining the features implicated in it, Biber interprets the dimension as being concerned with narrative *vs.* other types of discourse. He then calculates the scores of the different registers upon it. Registers such as romantic fiction, mystery, and adventure fiction obtained high positive scores on Dimension 2 while official documents, academic prose, and broadcasts scored negatively upon it. This distribution of registers, with fiction scoring high and more 'information based' registers scoring low, was taken as an indication of its validity; it effectively distinguished registers with a high concentration of narrative discourse from those with low levels of it.

Biber mentions two types of relevant but mutually dependent studies that must be considered when attempting research of this nature: microscopic analyses and macroscopic analyses. A microscopic analysis attempts an in-depth functional analysis of specific linguistic features while a macroscopic analysis summarises the overriding dimensions of linguistic variation in a language. In other words, microscopic studies examine the functions of specific grammatical constructions, such as time adverbials (ex. Chafe 1984a), while macroscopic studies provide a means of dealing with the similarities and differences between specific texts and registers. In some ways, Biber's MD method is a type of quantitative macroscopic study, but microscopic analyses are enlisted to help interpret the dimensions resulting from the factor analysis. Only by understanding, for example, what past tense verbs, third person pronouns, and public verbs have in common can one make a start at interpreting the factor that they help to compose (i.e. Dimension 2: "Narrative *vs.* Non-narrative Concerns").

Of Biber's findings, perhaps the most important is that there seem to be no clear-cut differences between speech and writing in English *per se* and that, in order to characterise the differences between registers, one has to appeal to a multi-dimensional model. Sixty-seven linguistic features are grouped into six dimensions in his analysis: Involved *vs.* Informational Production; Narrative *vs.* Non-narrative Concerns; Explicit *vs.* Situation-dependent Reference; Overt Expression of Persuasion; Abstract *vs.* Non-abstract Information; On-line Informational Elaboration. Clearly, some of these concepts hark back to earlier research on speech and writing. However, although significant differences were obtained between spoken and written registers along dimensions 1, 3, and 5, Biber states that even in these cases there was still substantial overlap.

Multi-dimensional research on languages other than English have obtained similar dimensions, interpreted based on the linguistic features associated with each factor. Biber (1995) speculates about the universal dimensions of register variation. Amongst his conclusions are the following:

1. Most languages seem to have an oral/literate dimension involving interactiveness and the elaboration of information (Type A). Features loading on the 'oral pole' include: 1st and 2nd person PNs; questions; imperatives; reduced features (contractions, fragmentation); stance marking features (hedges, emphatics, etc.). Features loading on the 'literate pole' include: nouns, adjectives, and dependent clauses. Conversation typically loads positively on the dimension while informational and expository prose load negatively.

2. Most languages have an oral/literate dimension dealing with production circumstances (Type B) but the defining features tend to differ for each language. Reportage of events in progress typically shows up in this dimension, as it is a register maximally constrained by production time.

3. Most languages have a dimension involving personal stance (Type C). This is not always robustly related to mode, but is frequently associated with interactiveness and personal involvement. It is marked by features such as hedges, downtoners, emphatics, and attitudinal expressions. Positively loading registers typically include public speeches, professional letters, and editorials.

4. Languages usually have oral/literate dimensions that are idiosyncratic. In English, for example, passivisation is a marker of abstract *vs.* non-abstract styles. Biber says: "These dimensions... show that in some cases, relatively idiosyncratic communicative functions particular to a culture and language are important enough to be reflected in underlying linguistic dimensions" (p. 253).

5. Aside from the oral/literate dimensions, all languages sampled have a dimension marking narrative registers. Features loading positively on the dimension include past-tense verbs, 3rd person PNs, temporal adverbs, temporal clauses, and special verb classes. Written fiction and folktales are the highest loading registers on this dimension.

6. Most languages also have dimensions involving the marking of persuasion and/or argumentation.

Jang's dissertation on Taiwanese (1998), completed after Biber's (1995) book, reported the following dimensions: Interactional *vs.* Informational Focus; Personal Expression of Emotion; Persuasion; Narrative Concern; Involved Exposition *vs.* Precise Reportage. Here, we can see the relations with the types Biber proposed:

Table 1: Comparison of text types in Jang (1998) and Biber (1995)

	Biber 1995
Interactional <i>vs.</i> Informational Focus	Type A
Personal Expression of Emotion	Type C
Persuasion	Persuasion/Argumentation Dimension
Narrative Concern	Jang 1998
Involved Exposition <i>vs.</i> Precise Reportage	Type B

Thus, there is a close correspondence between those dimensions reported as universal and those that Jang found in Taiwanese. Although the current study is not multi-dimensional in nature, the findings above will be salient when examining the differentiation of the Scottish Gaelic registers in terms of linguistic form. In particular, it will be seen that it is possible to relate many of the pivotal constructions comprising dimensions in other languages to functional equivalents in Scottish Gaelic.

2.4 Scottish Gaelic: Previous Register and Spoken Language Research

Although there is a dearth of empirical register work in Scottish Gaelic, it is useful to take stock of spoken language research in general in the language, which is more thorough although of a fundamentally different nature. The advent of portable audio recorders meant that the actual speech of human beings could finally be captured in natural settings and preserved for posterity. The first pioneers of this new technology in the Scottish *Gàidhealtachd* were folklorists such as John Lorne Campbell and Calum Maclean. The agenda in Gaelic spoken language studies has to a large extent always been dictated by urgency, as Ní Laoire (1993) calls it, the 'fire brigade approach'. In the case of Gaelic folklore research in the 40s and 50s, the last *seanchaidhean* – storytellers and caretakers of oral tradition – were old and their memories fading. Although waning for years, the Second World War saw the context for the transmission of oral culture, the house visit or *cèilidh*, largely supplanted by the wireless. In some cases, these individuals were trying to recall songs and stories that they had not performed for decades. Fortunately Campbell, MacLean and later field workers associated with the School of Scottish Studies in particular, managed to record thousands of hours of tape containing many songs, stories, interviews, and general lore that would have otherwise expired. Quick on the heels of this effort was the concentrated recording of moribund or endangered dialects, particularly those on the mainland and southern periphery. The *Survey of the Gaelic Dialects of Scotland* was begun effectively in 1951 under the direction of Kenneth Jackson at the University of Edinburgh (Ó Dochartaigh 1997). Part of the fieldwork methodology included collecting less-formal dialogues on tape where the discussions would be of a general nature, including talk about the participants' background and their recollections from youth. Although the morphological information is still being collated, the phonetic material has been published and researchers such as Ó Maolalaigh (1999) have demonstrated the value and research potential – particularly the diachronic implications – of this vast store.

Research in folklore and phonetics has accounted for most of the published work on spoken ScG. In relation to the number of studies in these areas, very little work has been done on register variation or examining the differences between Gaelic speech and writing. Mackinnon's (1977) sociological study of Gaelic usage contributes much to the consideration of register range, but Macaulay's (1982a) paper was the first to explore variation as such. He uses three short recollected dialogues to illustrate some lexical and code-switching elements in Gaelic buying and selling language. He also sets the scene for future work on register in Gaelic, providing a useful overview of the subject, particularly in terms of the formal/informal dimension and its implications for the case system, address forms and borrowing. He indicates that address forms (*thu* vs. *sibh*) vary along lines of status and formality, but that there are other variables such as familiarity implicated which do not allow simple characterisation of these as discourse markers. He concludes that more formal registers, such as religious speech and oral literature, tend to resist English borrowing and contain certain recognisable formulae (register markers). The main drawback with this paper is that he does not use transcriptions of actual speech; we cannot be sure whether his recollections are true to detail.

Ní Laoire (1988) examines an existing corpus of Irish conversations made by the folklorist Hartman for features of register variation. The conversations were mainly between a chairman (Hartman) and two participants who sat around a table with other listeners, drinking Guinness and tea and talking about anything they found interesting. However, she goes on to say: "None of the items which I shall outline below as variables which might be indicators of stages along the formality-intimacy continuum are striking features of the transcripts provided by Hartman" (p. 294). She then goes on to provide the items but does not indicate the sources of her observations. Like Macaulay's (1982a) paper, Ní Laoire sets the scene for future work rather than offering conclusive evidence. The main linguistic features of register variation that she highlights are: syntactic reduction,

such as loss of genitive case marking; use of older forms, e.g. synthetic verbs, in more formal registers; the incidence of more English borrowing in lower registers; interference from written forms in speech and vice versa; and possible borrowing of forms from a more prestige dialect, such as Munster, into less prestigious dialects. She points out that one could look at many more possible correlates of register than she mentions, in particular word order and syntactic variables.

Nancy Dorian has done the most empirical work on synchronic variation in Gaelic, but, again, most of it has been within the context of dialectal and language death work – particularly concerning East Sutherland Gaelic (ESG) – rather than examining register *per se*. In fact, she tells us (1994b) that she never set out to investigate it directly, but found some of her narrative data pointing towards certain correlates of stylistic variation along the formal and informal axis. She asserts that register variation in ESG interacts with a state of limited diglossia, with higher registers such as religious discourse and public speaking entailing the adoption of more prestigious, western dialectal forms in preference to lower prestige ESG forms. She mentions in another paper (1994a) that obvious loans and code switching from English, as well as profanity and strong interjections, are eschewed when speakers navigate registers that are more formal.

Folklorist John Shaw's (1999) paper is notable in that he attempts to uncover divisions within the Gaelic language for different speech styles. He finds that the root *beul* tends to indicate generally sound, time, and space while *cainnt* refers to "a stream of speech", a capacity to speak, the marking of people's origins by their speech, and expressive language. He examines some well-established traditional genres in Gaelic, such as mocking, satire, and praising and mentions that they have largely disappeared. Shaw also considers the special high register greeting language found in oral literature, used for effect rather than content's sake when characters are in the presence of royalty and the supernatural.

Finally, some researchers have noted that Biblical language can show up in higher register discourse, particularly formal written language (Meek 1990; Macaulay 1992). However, as Macaulay remarks, illiterate speakers will not have direct access to it. Moreover, although Meek points out that the retreat of the church in Highland communities has meant that younger Gaels are no longer getting a chance to absorb certain higher register codes, one wonders if this type of discussion is more typical of the Protestant rather than the Catholic religion.

Overall, there has been almost no empirical work on Gaelic register variation. What comments there are on the subject have generally been garnered as the result of introspection and informal observation, or as a by-product of dialectal research. To summarise, researchers have maintained the following about register variation in Scottish Gaelic:

<i>Higher registers in Gaelic are associated with:</i>	<i>Author(s)</i>
Less profanity or strong interjections	Dorian 1994a
Less syntactic reduction	Ni Laoire 1988; see also Macaulay 1996
Recognisable formulae	Macaulay 1982a
Resisting of borrowing from English	Macaulay 1982a; Ni Laoire 1988; Dorian 1994a
Some influence from Biblical discourse	Meek 1990; Macaulay 1992
Special greeting language	Shaw 1999
Switching to more prestige dialectal forms	Ni Laoire 1988; Dorian 1994b
Use of older forms	Ni Laoire 1988

Clearly, the extent of current knowledge concerning Gaelic register variation in Gaelic is limited and confined only to the dimension of formality. Dressler (1988) reviews a number of morphosyntactic effects of language decline and asserts that one of these is a tendency

towards mono-stylism. He says that dying languages tend to be used in gradually more casual settings until they are inappropriate for high status domains. One might ask if the reason that researchers have not located any other instances of variation in Gaelic beyond formal and informal is that the language has retreated to the point that there are only two broad contexts left to determine discourse level. However, Dorian (1994a) argues that register differentiation can persist even in languages that are mono-contextual and that it ought to be measured more finely than Dressler seems willing to do. Dorian anticipates that register variation can be witnessed to persist, when texts are collected along finer characterisations of context, wherever the potential for syntactic variation does. This is the thrust of the current study. In fact, when studying register variation in an endangered language, one must account for many more variables influencing linguistic form than in the case of a monoglot language situation. Internal language change, superstratum influence, differing proficiency levels amongst speakers (see Dorian 1981) and contextual vicissitudes coalesce to present a rich and complex web of linguistic variation.

2.5 Summary

Within the discipline of linguistics, echoing academia at large, there has been an entrenched habit of bestowing more esteem upon literate forms of language than speech, especially speech produced on-line in informal circumstances. In the 20th century, this position was redressed on a conceptual level as linguists generally acknowledged the primacy of speech in human language. In practice, however, while espousing this position, much syntactic research has continued to concern itself with structures characteristic of carefully edited prose rather than attending to the different, but no less valid intricacies of spoken language.

In the past thirty years, some linguists, especially those of a descriptive/functional inclination, have begun to empirically probe the ways in which speech and writing diverge from one another. This line of inquiry gained vitality from the advent of sophisticated computerised corpora in the 1980s, which has enabled researchers to compare language samples with unprecedented speed and rigor. Chafe's writings in particular have been influential, where he claims that typical speech tends to be involved and fragmented while writing is characteristically detached and integrated. In the early 1980s, empirical work on register variation as such had not yet come to fruition and studies were usually limited in both size and scope, often comparing only two language varieties with one another. Controversies surfaced, one of the most notable being whether speech or writing is more syntactically complex. Some researchers, such as Halliday, advocated for speech, while most others, such as Chafe and Thompson, found writing to be more syntactically intricate. This debate, along with that relating to the distribution and conceptual basis of subordination, highlighted the importance of coming to a consensus on definitions and methodological issues. Researchers questioned the ramifications of selecting some registers over others as representative of speech or writing and a dialogue ensued over the impact of language producers on linguistic form. At this point, researchers were still concerned with differences between the modes (speech and writing), as opposed to variation between different registers.

The work of Biber and a small number of others adopting his methodology has managed to rectify many of the limitations and aberrations of previous work on the linguistic divergence of speech and writing. He included a wide variety of different registers in his sample, thus permitting the expansion of his conclusions beyond strictly writing or speaking. He also counted a greater number of linguistic features than had been done up to that point, allowing a finer characterisation of the differences between the texts. Finally, by employing multivariate statistics, especially factor analysis, he was able to distil a number of global contrasts between the registers from what would otherwise have been an overwhelming mass of data.

Biber's most important conclusion was that there were few distinct differences between speech and writing *per se* and that it was necessary to appeal to a factorial model to account for all of the variation in the data set. Incorporating the results of multi-dimensional studies on other languages, Biber has proposed a number of 'universal' dimensions of register variation. Later studies, such as those conducted by Greenbaum, Nelson, and Weitzman, have generally confirmed Biber's results, revealing more variation within the modes than between them. However, spontaneous spoken language, such as informal conversation, has been found in numerous studies to be unique in many ways. Miller and Weinert's (1998) in-depth study of SSL builds upon the prior research of Biber, Chafe, Halliday, Ochs and others to survey the form and function of a wide range of syntactic and discourse-related features characteristic of on-line speech. Unlike most previous studies, they incorporate data originating from language producers outwith the university population and a wide number of languages in addition to English. They assert the importance of their findings not only for register variation research, but also for the disciplines of theoretical syntax, language acquisition, language typology, and historical linguistics.

Little empirical work has been done on either register variation in the Goidelic languages or the syntax of their spoken forms. The small amount that has been written on the subject has been arrived at largely through introspection or as a by-product of other lines of investigation such as dialectology and language death. Although Dorian has countered Dresser's position that dying languages, such as Scottish Gaelic, do not evince register variation on theoretical grounds, the empirical proof needed to settle the question has not yet been advanced. One of the aims of the present study, grounded in the work of the above authors (especially Biber and Miller and Weinert), is to move towards supplying this proof.

3 Scottish Gaelic Sociolinguistics

The following chapter briefly surveys a number of topics relevant to the sociolinguistic study of Scottish Gaelic: history and dialectology; the economy and demography of its speakers; characteristics of use and register range; and the patterns of its decline.

Map 1: Traditional Gaelic-speaking Scotland



3.1 Brief History and Dialectology

Scottish Gaelic, or *Gàidhlig* [gailij] as it is referred to by its speakers, is a member of the Celtic family of Indo-European languages and, furthermore, part of the Goidelic branch of that family along with Irish and Manx Gaelic. While there was once a good deal of argument concerning ScG's origins in Scotland, the strong evidence for its importation by Irish immigrants has all but stamped out an earlier 'nativist' stance. Although it is not clear exactly when or where Irish-speaking people originally landed in the territory known today as Scotland, it is assumed that this coincided with the decline of Roman power in Britain and was part of a larger infiltration of Britain's west coast, about the late 4th century AD (Macaulay 1992). The Gaelic languages of Ireland and Scotland are today not mutually intelligible without considerable priming, but there was close cultural, material and linguistic exchange between the two countries until the collapse of the old Gaelic orders in the 17th (Ireland) and 18th (Scotland) centuries.

To fully appreciate the proliferation of the Scottish Gaelic dialects, it is best, of course, to refer to the greater Goidelic continuum, which at one time would have stretched from the southern-most tip of Munster in Ireland up to the Butt of Lewis in Scotland and included the Isle of Man in its south-eastern periphery. However, as the language has retracted and eventually disappeared from many areas along this geographical spread, the linguistic gulf in the remaining regions has been perceived to widen. Although we now speak of the 'languages' of Irish, Scottish Gaelic, and Manx, this is due in part to the demise of the dialects that once acted as bridges between these areas, e.g. that of Rathlin Island.

In an oft-cited passage, Jackson (1968) says that the Scottish Gaelic dialects divide, in many ways, roughly east-west and north-south but that perhaps the best way to categorise them is as a 'central dialect' and a 'peripheral' group of dialects. The central dialect is characterised as largely innovating and homogenous. The peripheral group is conceived as more heterogeneous and fragmented but it has certain broad features in common, notably a tendency to preserve archaic forms of the language:

... it is possible to say in very broad terms ... the central dialect covers the Hebrides as far south as Mull and sometimes further, Ross exclusive of the north-east corner, Assynt, Inverness-shire, western Perthshire, and mainland Argyll roughly north of Loch Awe ... the peripheral dialects comprise Caithness and Sutherland exclusive of Assynt, the north-east corner of Ross, Braemar, eastern Perthshire, the rest of mainland Argyll with Kintyre, and Arran. Moray and the adjacent lower region of the Spey, the wide valley of Strathspey from Rothiemurchus to the Moray border, may go with the peripheral dialects, linking up with Braemar and east Perth. (pp. 67–8)

Thus the Gaelic of the Hebridean region as a whole largely goes with the 'central' dialect while that of the mainland and more outlying areas, where it is now generally moribund or non-extant, tends towards the characteristics of the 'peripheral' group. Discussions of the differences between a 'central' and 'peripheral' prototype have tended to centre around the following key features:

- Pre-aspiration before the stops *-c*, *-p*, and *-t*: absent in the peripheral group, of different varieties in the central dialect (Gleasure 1994), e.g. *cat* 'cat' – Uist /ka^ht/, Arran /ket/
- Svarabhakti: where vocalic epenthesis occurs, it tends to be somewhat stressed and 'clear' (resounding the previous vowel) in the central dialect, while peripheral dialects tend towards the older schwa [ə], or [i] (Gleasure 1994), e.g. *carbad* 'wagon' – Uist [karabət], *vs.* Arran [karsbət]
- Diphthongisation of some stressed vowels before the sonorants *ll*, *nn*, and *m* where they are not followed by a vowel: periphery tends towards preserving older non-diphthongised pronunciation (Gleasure 1994), e.g. *mall* 'slow' – Uist [maul] *vs.* Arran [maL']

Map 2: The Southern Isles of the Outer Hebrides



- Breaking of Old Irish long *é*: central areas have tended towards diphthong pronunciation while peripheral areas have preserved older pronunciation (Jackson 1968), e.g. *beul* 'mouth' – Uist [bial] *vs.* Arran [be:l]

Within the Outer Hebrides, it is possible to make a division between the Gaelic of Lewis and all islands south of it (Harris, N. Uist, Benbecula, S. Uist, and Barra) such that Borgström (1940: 228) refers to the “‘northern dialects’ of Lewis” and “the Southern Hebridean dialects”. There are isoglosses such as *bùrn* ‘water’ in Lewis and *uisge* elsewhere, lexical and as well as significant phonological variation (Borgström 1940: 228). Gaelic speakers themselves, on both sides, are acutely aware of these differences. Subdivisions can be further specified between the islands themselves (e.g. Catholic *vs.* Protestant Uist) and between bordering villages. With such rich phonological variation present in their speech, Hebrideans are sensitive even to the slightest shibboleth¹⁴ and are often able to place each other with remarkable accuracy.

Scholars have long lamented the unavailability of dialectal data on the language, and it has been customary to proceed with caution when making assertions about the area. The publication of the phonetic sections of the *Survey of the Gaelic Dialects of Scotland* (Ó Dochartaigh ed. 1994–7) has finally made it possible to begin to rectify this situation.

The grammatical sketch that is presented in Appendix 1 is meant to be applicable to most of the surviving Gaelic dialects although there is a certain bias towards the Gaelic of the Outer Hebrides and especially that of North Uist, where most of the fieldwork was done for this research. For more information on Scottish Gaelic dialectology, see Ó Dochartaigh (ed. 1994–7: vol. 1) and the references therein.

3.2 Demography and Ethography

At the end of the 11th Century AD, Gaelic was the common language of nearly all parts of the Scottish mainland.¹⁵ Today, however, the Gaelic-speaking region or Gàidhealtachd is confined generally to the Highlands and the islands off the west coast of the country. The last reported Census (2001) revealed that approximately 45% of Gaelic speakers (out of a total population of 58,552) now live outside of this area, with concentrations especially in the larger cities. Yet the vast majority of these urban Gaels are emigrants from traditional Gaelic speaking areas and there are low rates of second-generation transmission outside these areas. (The following statistics are from MacKinnon 1995, which analysed the 1991 Census data, but the trends are still applicable.¹⁶) In the Outer Hebrides, 80.7% of families (N=1310) with two Gaelic speaking parents raised children who spoke the language. In the Skye and Lochalsh area, this figure was slightly higher, at 87.1% (N=190 families). For the rest of Scotland, 60% of these families (N=730) had children that were Gaelic speaking. The most dramatic difference occurred between two parent families where only one was a Gaelic speaker. In the Outer Hebrides, 24% (N=610) had Gaelic-speaking children. The figure was again higher for Skye and Lochalsh at 40.4%. In the rest of Scotland, out of a total of 2460 families of this type (the highest proportion of the population of Gaelic speaking families in this area), only 6.8% passed the language on to their children.

Although Gaelic speakers are still, though decreasingly, to be found throughout the Highlands and Islands as a whole, the 2001 Census showed that it is only in the Outer Hebrides (i.e., Barra, the Uists, Harris, Lewis) and Northern Skye (Kilmuir) that Gaelic speakers constitute a majority of the population: 59% and 57% respectively.¹⁷ The population density in the Western Isles region is very low overall with most of the

¹⁴ For instance, Cairinish, N. Uist, joins with Benbecula and South Uist in having *fhìn* as its 1S reflexive form whereas most of N. Uist uses *fhèin*. When teased by their contemporaries about this, Cairinish children used to rhyme in defence: *cha leig thu leas a bhih a' gaireachdainn – 's ann a thu mi fhìn à Cairinis* 'you don't have to laugh – I myself am from Cairinish'.

¹⁵ For histories of the language see: Durkacz (1983); MacInnes (1992); MacKinnon (1974); and Withers (1984).

¹⁶ The author expresses his appreciation to Dr Wilson McLeod for bringing this paper to his attention.

¹⁷ Tiree used to be able to claim the same, but the number of Gaelic speakers has fallen from 58% to 49% since 1991.

settlements in the form of scattered crofts. What towns there are, Stornoway on Lewis and Balivanich in Benbecula, tend to be the most English-dominant areas of the islands.

3.3 Economy and Community

Many of the inhabitants of the Outer Hebrides living outside the towns are involved in crofting to some extent but few are full-time crofters. Often, families keep a small number of sheep or cattle, supplementing their main income with the proceeds of carcasses or wool. In the late 20th century, however, sheep became a near liability. Their market value is still so low that they are now often kept out of loyalty to a way of life rather than economic profit. Although prices have come up slightly in the last few years, the BSE scare and foot-and-mouth epidemic have had a severe impact. If the agricultural grants that crofters rely upon were to disappear, crofting would cease to be a viable occupation for most. Fishing is the other main traditional occupation of the islands. There are more full-time fishermen than crofters at this point, but the fishing industry, again, is not as fruitful as it once was. Tourism has become increasingly important to the economic health of the Outer Hebrides and many people earn a substantial income from house letting or bed and breakfast. Telecommunications and computerisation have also been growth areas and are seen by some as solutions to the problems inherent to industrial development in the area, among which are its remoteness and lack of pre-existing infrastructure.

Although the Outer Hebrides are one of the poorest regions in Britain, outright poverty is scarce or at least camouflaged by consistently modest housing and a general lack of the unsightliness that typifies urban poverty. There is a strong tradition of caring for the older members of the community and the less well-off, seen in neighbours fostering individuals or families, providing such things as a modicum of fuel and food for them. Although less common than in the past, there is still much co-operation within communities, especially at certain crucial periods of the year, such as at shearing and lambing.

Perhaps the most recent major blow to the islands was World War II, which practically depopulated some areas of their young male generation. The sadness that ensued from its casualties, horrors and losses, and the new, more technological and material world that it ushered in, are often cited as major factors in the breakdown of the traditional culture, in particular, the practice of house visiting or the *ceilidh* as it is called in Gaelic (see MacDonald 1999). This was the setting for community joining, entertainment and the transmission of Gaelic culture in the form of songs, stories, music and dance. In around the mid 20th century, the word was semantically extended to accommodate a new form of visit – to the community hall – for an evening of group dancing or, with the modifier 'traditional' attached to it, a concert of Gaelic singing, instrumental music, and Highland or step dance. The major difference, however, is that older *ceilidhs* would have seen many more participants doing their own 'party piece'; a clear effect of the passivisation resulting from two generations of dependence upon electronic media.

Compared to the rest of Scotland, the Outer Hebrides, in particular the southern isles (Uist and Barra), have far fewer professionals in their workforce (MacDonald 2000). For the young adult who graduates from college or university, there is scant opportunity for returning to a job commensurate with his or her educational level. The Hebrides are also an area with a growing senior population and diminishing youth. This fact has clear implications for the Gaelic language as older, Gaelic dominant speakers are replaced by a fewer number of young, English-dominant speakers. Depopulation has also been a great problem and the tabloids warn of a future crisis if current rates continue.

Religion, traditionally one of the vanguards of Gaelic usage, is still influential in the lives of the people in the Outer Hebrides¹⁸ although, like the rest of Scotland, it is losing its sway with the younger generation. The dividing line between Protestantism and Catholicism is Benbecula, with most of those north of the island subscribing to the former, and those south, the latter. On Benbecula itself, there is a mix of the two

¹⁸ See the fervent debate covered in the *West Highland Free Press* and the *Stornoway Gazette* concerning ferry sailings and flights to and from Lewis on the Sabbath.

religions. Gaelic-medium congregations are still to be found but, inevitably, their viability is tenuous in most areas (MacKinnon 1998).

One of the main reasons cited by Gaels for the withdrawal of the language is the influx of people from outwith the *Gàidhealtachd* and the concomitant decrease in the likelihood that all members of any given conversation will be Gaelic-speaking. Incomers are sometimes resented for their perceived diluting of the language's strength on the islands. Dorian's work in Mull (1981b) reported concomitant feelings of resentment amongst non-Gaelic speakers when Gaelic was used to exclude them from conversations. However, it is surprising to many incomers that Gaels do not simply speak Gaelic in their presence despite them, especially when they are not actively participating in a conversation. Some Gaels respond to this by saying that they have had it instilled in them to always be courteous, with this kind of exclusion considered rude. But it also betrays a pervasive and deep insecurity regarding the language and the belief that it is not worth learning if it is not one's mother tongue.

3.4 Language Usage

3.4.1 Age, Occupation, and Gender

Today, the only Gaelic monoglots are very young children whose parents have made an active policy of banning English in the house. However, when one begins to look in earnest for these families, one realises how few there are. In 1998, as part of a funded research project that the author was involved in, a search was made in the Uists for L1 (not necessarily monoglot) Gaelic-speaking children between the ages of three and five. Out of a population of nearly 6,000 only 20 of these children could be located. In N. Uist, in 2006, there were about six pre-school children¹⁹ who spoke Gaelic in the home, out of approximately 50 in total. To make matters worse, there is currently no full-time Gaelic nursery to be found in N. Uist (S. Uist has one), so working parents without Gaelic who wish their children to have the language do not have any viable options.

Even in cases where children are, for all intents and purposes, monoglots, it is only a matter of time before their exposure to media, school, and non-Gaelic speaking children result in rapid English acquisition. Many Island parents begin child-rearing hoping to have a Gaelic household, which does not seem too exotic as it would have been their experience when young. However, once their children begin making steady contact with the outside world, English almost invariably takes over in a predictable pattern. Increasingly, children begin answering back in English until parents give in and Gaelic gets consigned to imperatives and choice phrases rather than being a conversational medium. In the 1991 Census, 49.5% of children aged 3–15 were Gaelic-speaking (although certainly some of these would be 'semi-fluent' speakers, such as Dorian (1981a) located in her study of East Sutherland Gaelic). Showing the decline that ten years can bring on however, only one fourth of the children who entered island schools in 2000 were Gaelic speaking (MacDonald 2000).

MacKinnon (1998) compared the Gaelic-English usage patterns found in a 1994–5 survey of Gaelic speakers throughout Scotland (Euromosaic Project) and a similar one completed in 1988 (Language Maintenance and Viability Survey). He reports that Gaelic usage was being maintained to a certain extent in the pub and between neighbours but that it was falling in shops and in the church. The intergenerational use of Gaelic – between parents and children – was seen to have contracted from the levels of 10 years prior: the number of respondents indicating that they 'always' or 'mainly' used Gaelic with their children was minimal. As expected, this finding was echoed for the use of Gaelic between siblings themselves. The rise in the number of children speaking only English with each other is particularly striking, being 10% higher in 1995 than approximately eight years prior (tables from p. 61):

¹⁹ Two of these are our own children, aged 4 and 2 at the time.

Table 2: Language-maintenance and viability survey 1986–8 (Western Isles Gaelic speakers: N=224)

Language Usage	By subject to grand-parents	Between subject's own parents	Between subject and spouse ²⁰	By subject to own children	Between subject's own children
	No. (%)	No. (%)	No. (%)	No. (%)	No. (%)
Gaelic	212 (95.9)	212 (96.4)	117 (72.7)	80 (55.2)	45 (35.5)
Both equal	3 (1.4)	1 (0.4)	18 (11.2)	37 (25.5)	29 (22.8)
English	6 (2.7)	7 (3.2)	26 (16.1)	28 (19.3)	53 (41.7)
Total	221 (100)	220 (100)	161 (100)	145 (100)	127 (100)
N/R or N/A	3	4	63	79	97

Table 3: Euromosaic survey 1994–5 (Western Isles Gaelic speakers: N=130)

Language Usage	By subject to grand-parents	Between subject's own parents	Between subject and spouse	By subject to own children	Between subject's own children
	No (%)	No (%)	No (%)	No (%)	No. (%)
Gaelic	195 (88.2)	107 (83.6)	45 (59.2)	38 (49.4)	18 (28.6)
Both equal	12 (5.5)	7 (5.5)	12 (15.8)	15 (19.4)	13 (20.6)
English	14 (6.3)	14 (10.9)	19 (25.0)	24 (31.2)	32 (50.7)
Total	221 (100)	128 (100)	76 (100)	77 (100)	63 (100)
N/R or N/A	39	2	54	53	67

If we consider Gaelic-speaking parents in Scotland at large sampled for the 1994–5 study, approximately eight percent reported their children as speaking 'only' or 'mainly' Gaelic (MacKinnon 1998: 63).²¹ Clearly, the Outer Hebrides – along with Skye and Lochalsh – still exhibits higher levels of language transmission than the rest of Scotland, but there is no guarantee that this will be the case for much longer. MacKinnon states: "Unless tackled promptly and effectively, the prospect of Gaelic continuing as a family and a community language is likely to cease with the present generation" (p. 55).

In an earlier study MacKinnon (1994) found that crofters, after professionals and housewives, evinced the highest intergenerational Gaelic usage levels. We can expect the decline of crofting to have ramifications on the continuation of the language. He says: 'There were very distinct and significant differences in Gaelic usage levels in original and

²⁰ There must have been a typographical error in the original publication where this column duplicates the heading "By subject to own children". However, on page 55, MacKinnon says that they surveyed use: "...from the generations of the subjects' grandparents and parents, to exchanges between spouses/parents of the present generation, by parents to children, and between children themselves [italics added for emphasis]." Clearly the heading was meant to be "Between subject and spouse".

²¹ There is a typographical error on this page as well, a duplicated number, but the frequency for children only speaking English with each other can be obtained from adding up the rest of the row and subtracting from 100.

present-day families, and in intergenerational Gaelic maintenance, between the respondents who were members of a crofting family and those who were not' (p. 125). Examining the 1981 Census data, MacKinnon found there were still communities which had a proportion of under 25 year-olds reporting Gaelic usage levels as high or higher than individuals older than that. He says: 'The main bulwark of their linguistic and cultural integrity might seem to lie in their essential character as crofting communities' (p. 126).

Apparently, there is no empirical research indicating how age correlates with language proficiency. However, although schooling in the past century has ensured that virtually all adult Gaelic speakers are capable of communicating in English, many older speakers are still more comfortable speaking Gaelic. At the risk of over-simplification it seems in Uist that there is a band of people, generally over 65 years of age, who are more Gaelic dominant. There is another younger band who are equally comfortable in both languages. Speakers less than about 45 years old may be functionally fluent in Gaelic but tend to be English dominant. In the crofting townships of North Uist, the generation born in the late 1960's is reckoned to be the last for whom Gaelic would have been the language of the playground. The last school in Uist with children who naturally spoke Gaelic to one another was Stoneybridge primary, in South Uist, where this continued until the 1980's at least. The coming of television is the attributed cause of the switch. For Gaelic speakers below the age of 30 or so – the 'television generation' – the language which they use with their contemporaries is almost exclusively English. One finds, to varying extents, that they will only use Gaelic in restricted situations, such as being paid to talk on the radio or the television, and/or with only the eldest members of the population. The only situation in which the author has heard Gaelic used as a conversational medium within this age group is at the pub, and only after a considerable amount of 'tongue-loosening'. Whenever we interpret Census data, the figures must be considered in terms of a substantial proportion of returns who, regardless of their proficiency in the language, contribute a negligible proportion of overall Gaelic output.

There has been some evidence that young women in Gaelic-speaking communities evince less support for the language than men and women of other age groups (MacKinnon 1977; 1994; 2000). Although, overall, older women reported the highest usage levels in MacKinnon's 1994 report, young mothers reported the lowest. This has clear implications for the transmission of the language. Young women were also reported as being more likely to migrate than young men and, as was seen above, Gaelic usage is lower overall amongst children on the Scottish mainland. MacKinnon stresses that these findings should not be taken as indicating lower *loyalty* amongst young women towards their first language; those who do move away to the mainland are often the most fervent Gaelic supporters.

Constantinidou (1994) explores gender effects in the contraction of East Sutherland Gaelic, citing research by MacDonald (1987) *et al.* indicating that women are more likely than men to code switch when in the presence of non-speakers. She also identifies a difference in the stories of women and men concerning situations where Gaelic was spoken around perceived non-speakers. Women brought up occasions when they were 'caught out' by people who actually did have the language, and thus were interpreted as more sensitive, therefore, to the faux pas they had committed. In contrast, men's stories lacked this social self-consciousness; there was no mention of being caught, only of cohesion between the speakers. She states: "In the female world Gaelic contexts of privacy could be breached, whereas in the male world privacy could be achieved and maintained by means of Gaelic" (p. 121).

3.4.2 Register Range and Bilingualism

3.4.2.1 Spoken Registers

Although Gaelic was once the dominant tongue in Scotland and used at all levels of situational formality,²² it is today most often the medium of a narrow band of registers. As virtually all Gaelic speakers are fluent in English, they have a choice as to which language, or proportion thereof, they use in a given situation. Macaulay (1992) says that there is a 'normal' range of register differentiation in Scottish Gaelic. However, the more formal, official, technical, abstruse or field-specific the subject matter, the less likely it is that Gaelic will be the language of choice (Macaulay 1982a; Thomson 1994). Exceptions to this tendency are to be found in religious discourse or technical discussions involving crofting or fishing, but there are no spoken registers which one can now guarantee will be Gaelic-medium as so many other sociolinguistic variables assert themselves in determining language choice.

Macaulay (1982a) mentions that it is easier to imagine contexts where Gaelic is likely to be excluded than the converse. Perhaps the most reliable context for primarily Gaelic usage would be one where the conversational setting and theme are informal and non-field-specific, and where the participants are all older, native speakers, familiar with each other, and in a domestic setting. As soon as one steps out of the controlled environment of one's house, the likelihood increases that some participants will be English-monoglots or semi-speakers.²³ Even in such an idealised situation, however, code switching, or at least borrowing, is the rule rather than the exception. This is partly due to the lack of a large magnavocabulary in Gaelic, especially for terms associated with the modern world.²⁴ However, when considering the association of register and language choice in spoken Gaelic, perhaps the crucial determining factor is not so much the formality or informality of the subject matter, but rather *who* the participants are and *where* they are. The author's experience has been that individuals who normally speak Gaelic with one another tend to do so regardless of the topics at hand. Dorian (1986) concurs, saying that her sample of ESG speakers used Gaelic for the communicative situations in which individuals spend most of their time, as well as for discussing abstruse topics:

They could narrate, argue, joke, gossip, tease, discuss health, community, and national affairs, and business matters; they were fully able to exchange news, advice, plans, and opinions on matters grave and trivial... Because of the ease with which English loanwords are taken into ESG... even the most up-to-date and technical subjects could be discussed in Gaelic, and were. (p. 259)

There is reason to suspect that Gaelic speakers do not switch to English when discussing formal topics because of the theme *per se*, but rather because there is a correlation between certain types of discourse and the presence of non-Gaelic speakers. If the conversation between two Gaels who normally speak Gaelic with each other switches to discussing the latest cloning experiment, they do not need to change language or grammar, but only the vocabulary they are using, and this is easily taken on loan.

Macaulay was clearly not only thinking of speech though and, as a case in point, the only biology publication – a field-specific domain – available in ScG is a school textbook by Ruairidh MacThòmais (1976). He says this about his effort:

... bha mi riamh gu làidir de'n bheachd gu bheil a' Ghàidhlig glè chomasach air rudan ùra a thoirt a-steach thuice fhèin agus gur h-e eachdraidh thruagh nan trì ceud bliadhna chaidh seachad a bha gar bacadh anns an dòigh seo ... Mar sin cha bu ruith ach leum gu dhol an sàs anns an obair seo.

²² For instance, there were once formal terminologies for linguistics, rhetoric, and law (Macaulay 1982a).

²³ It is not unusual to hear conversations between a fully fluent Gaelic speaker and a so-called 'semi-speaker' (often simply an unconfident speaker), with the first speaking entirely in Gaelic and the second, following every word, but speaking only in English.

²⁴ See McNeir (2001) for one government-sponsored attempt to rectify the situation.

Is dòcha gun deach an ruith gu dabhdail an ceann greise. Chan e obair fhurasda a tha ann an eadar-theangachadh de'n t-seòrsa seo ... Tha deagh fhios agam gum bi pàirt de'n eadar-theangachadh doirbh, is gum bi grunn de na facail ùra a dheilbh mi trom no mi-fhreagarrach an dòigh eile. Thig atharrachadh an sud 's an seo mar a thig luchd-sgrìobhaidh eile chun a' chuspair seo, agus gu cuspairean eile tha an dlùth-cheangal ris. Ach tha mi'n dòchas gu bheil toiseach fàis againn an seo, agus gum faic sinn iomadh cuspair eile ris nach robh dùil, air a làimhseachadh tre Ghàidhlig.

Translation (by current author): I have always been of the opinion that Gaelic is very capable of embracing new things and it is [only] the sad history of the last three hundred years which has been hindering us. As such, I rushed at the chance to get involved in this work.

Perhaps my initial rush turned to hesitation after a while. Translation of this kind is not easy work. I know well that part of the translation will be difficult [to understand], and that many of the new words that I created will be heavy or infelicitous in another way. Change will occur as other writers delve into the subject and other subjects that are closely related to it. But I hope that we have a burgeoning growth here and that we will see many other subjects that we never expected dealt with through the medium of Gaelic.

The issue of magna-vocabulary is a pervasive one. MacThòmais's assertion that Gaelic is capable of taking on new terms is certainly true, but the litmus test is whether the public at large adopts expressions once created and, in general, they do not. If the linguistic patrons of world languages such as Russian and French struggle to maintain purity in the face of global Anglicisation, Gaelic would seem to have little hope in this regard. A dictionary of broadcasting and parliamentary terms has now been published (McNeir 2001) – including many new words for modern concepts in general. We can be sure that very few of these will enter common parlance, a situation that can be replicated in many other languages around the world. However, the Gaelic radio service, *Radio nan Gàidheal*, is listened to by most Gaelic speakers and there is the potential that, over time, some of these may be accepted by some. Or at least used by them when speaking on the radio.

Due to the commitment of Gaelic radio and television personnel towards register expansion in the language (see Lamb 1999), there is now a wide range of media output found in Gaelic. The Gaelic radio (BBC) broadcasts on 103.7–104.3 FM, for roughly 65 hours per week, with greater output on weekdays (e.g. 7.30am–12pm; 2pm–7pm; 9pm–10.30pm) and reduced output on the weekends (about four hours on Saturday and one and a half on Sunday). Television broadcasting is more intermittent, with occasional children's programmes in the morning and afternoon, and regular adult broadcasting on Thursday evening. Many programmes are scheduled after 12am to fulfil broadcasting quotas, giving prime-time slots to more mainstream programming. A glance at one week's schedule of Gaelic broadcasting in the *Stornoway Gazette* (c. 2000) includes these types of programmes:

Television:

- News programmes (*Telefios*, *Telefios na Seachdainn*)
- Children's programming (*Gocam Go*, *Splaoid*, *Dotaman*, *Coineanaich Ghòrach*, etc.)
- Early teenager variety show (*Dè a-nis?*)
- A travel programme (*Air Chuairt*)
- A current affairs debate programme (*Cunntas*)
- A church service (*Glòraich Ainm*)
- A soap opera (*Falach Fead*)
- An adult variety show – looking at famous people (*Sin Thu Fhèin*)

Radio:

- Morning news programme (*Aithris na Maidne*)
- Weekend summary news (*Aithris na Seachdainn*)
- Morning talk show (*Coinneach MacIomhair*)
- Music programmes (*Caithream Ciùil*, *Crunluath*, *Ceòl an Nos Ùr*, *Sruth na Maoile*, *Bàrd nan Òran*)
- Chat and music programme (*Mìre ri Mòir*)
- Religious and inspirational programmes (*Dèanamaid Adhradh*, *Tro Shàmhchair an Fheasgair*)
- Children's programme (*Aileag*)
- A dedication show (*Dùrachdan*)

Some of these are regular shows, but many come and go throughout the year. For example, among the types of televised programmes current to the early 21st Century are game shows, teenage variety shows (*Dè a-nis*), a current affairs programme (*Eòrpa*), a documentary on the Book of Kells, and programmes on the history of Scotland. The impression is that one is able to find a similar breadth of media registers in Gaelic compared to English broadcasting, but not nearly the same frequency (or quality, some would argue). This tendency, for breadth rather than depth, is similar to what one finds in the print media and written registers at large.

3.4.2.2 Written Registers

In the 2001 Census, there were 58,652 persons (cf. 65,978 in 1991) who returned as being able to speak Gaelic, 65,674 who could read and write the language and 92,396 who claimed to be able to understand it. Ability aside, only a minuscule proportion of those with Gaelic literacy would ever choose, or have reason, to write the language on a regular basis. Indeed, the most informal of written communications are normally in English, even between individuals who would normally choose to converse in Gaelic (Macaulay 1982a).²⁵ Much of this can be attributed to a lack, until the 1980s, of Gaelic-medium education and over 100 years of prohibitionist and awkward school policies have done great damage to the status of the language overall. Gaelic writers – that is, the small number of literati as such and those working in the media – have usually been through a university degree in Gaelic or Celtic, and in some cases now teach in these departments. This small, highly educated group is responsible for the majority of the prose, poetry, children's books, media output, etc., extant in the language.

In chicken-and-egg fashion, there is a lack of forums for Gaelic writing and a concomitant lack of incentive for Gaelic writers. In 1989, *Gairm*, the only periodical dedicated to Gaelic writing solely, had sales of about 1000–1200 per quarter and an estimated readership of 4000–5000. The *West Highland Free Press*, *Stornoway Gazette*, Highland editions of the *Press and Journal*, *Oban Times*, the *Inverness Courier* and the *Scotsman* all usually feature weekly Gaelic columns. *An Gàidheal Ùr*, a monthly Gaelic newspaper, is published with the *West Highland Free Press*. There is also a bilingual quarterly periodical, generally for and by learners of the language – *Cothrom* – published by CLÌ, which features a great variety of writing. Between these publications one can find a wide range of genres: reviews; editorials; letters to the editor; short stories; general interest stories, cooking columns; news articles; Gaelic learners' lessons; advertisements; children's corners; employment notices; and so on. The situation has improved for Gaelic writing noticeably in the past 15 years. Whether this has been accompanied by an improvement in Gaelic reading is another, unanswered question.

²⁵ Ronald Black of the University of Edinburgh said once (1998) that 'letter writing in Gaelic today is a deliberate antiquarian exercise'. However, as Dr Wilson McLeod later pointed out to the author, this comment may not be so valid today in the face of the growth of e-mail correspondence, which benefits from being more informal and 'oral'. It would be interesting to know if Gaels who might not have otherwise corresponded in Gaelic before may sometimes do so today using this new option, where orthography and grammar tend to be less important than the communication at hand.

Imaginative writing has been limited mainly to poetry and the short story; longer fiction was almost unknown in the language until the early 21st century, save for a few novellas by writers such as Iain Mac a' Ghobhainn (Iain Crichton Smith), Alasdair Caimbeul, Tormod Caimbeul, and Tormod MacIlleathain (Norman MacLean). The publishing house Clàr's *Ur-sgeul* series has meant that we finally have longer, modern fiction in print. Aonghas Pàdraig Caimbeul's *An Oidhche mus do Sheòl Sinn* (383 pages), won the Saltire Scottish Book of the Year Award in 2004 and certainly raised the ante in terms of what is deemed possible in Gaelic fiction. However, with such a small readership, there are meagre returns awaiting even the most well-received of would-be Gaelic novelists and one wonders about the long-term viability of these endeavours. The themes of Gaelic short stories and poetry tend to be inward looking, often framed by the bounds of the *Gàidhealtachd* itself. To read about Gaelic speakers in outer space or even conversing with a Edinburgh bus driver seems to require too much suspension of belief.²⁶ Gaelic drama (see Macinnes 1994 for an overview) has a long history, and some fine theatre groups such as 'Tosg' regularly tour Scotland. Due its being an oral art, at least in the delivery, it has been able to reach a greater proportion of Gaelic speakers than most print media.

Other types of writing, particularly non-fiction that touches on the wider world, have generally not been deemed viable since: Gaelic speakers tend to be more comfortable reading in English and they are readily available and perceived to be of a higher quality in English. As Thomson remarks (1989; cf. 1976): 'There is only one book on football, very little on politics, nothing on conservation, nothing on fish-farming, nuclear policy, space exploration, the Royal Family, fast food, feminism' (p. 40). There has been little in the way of writing that is expressly academic, but much that has been written by academics. Certainly, high-register, expository prose is possible to find in the language, often dealing with literature, folklore, and linguistics. But, as previously mentioned, there is a very small market or patronage for Gaelic-medium textbooks and academic periodicals. In earlier times, particularly pre-World War II (Meek 1990: 16), high-register prose was modelled upon the language of the Gaelic Bible. Since then, prose *per se* has moved towards greater colloquialism. (Even the Bible itself has been updated to reflect contemporary usage.) Finally, industrial and official documents found in Gaelic are often transparently tokenistic, especially when they are based upon more complete and concise English versions. When placed alongside original English sources, the extent of the inappropriateness of word choice and parasitism is readily observed (Cox 1998, McLeod 2000). Certainly there are documents of this sort written in idiomatic, edited prose, but it is not unusual to read writing issued by explicitly Gaelic organisations that is rife with orthographical mistakes, inconsistencies and errors as regards grammar and case marking, and 'something based on a dense English set of sub-clauses which just looks contorted in Gaelic' (Cox 1998: 81).

3.5 Summary

Scottish Gaelic was once spoken throughout much of the Scottish mainland, but is now confined mainly to its Western periphery and the Islands off the coast of this region. Language transmission in these areas has declined markedly in the past fifty years and research such as that conducted as part of the Western Isles Language Plan (2005) indicate that Gaelic will cease to be a community language in the next generation or two if current trends continue. The economy of most Gaelic speaking areas is based upon subsistence farming and fishing with fewer professionals in the workforce relative to other areas of the country. Overall, the remaining Gaelic-speaking areas are some of the poorest and most isolated regions in Britain. Part of the language's decline has been attributed to the increase of English monoglot incomers. As virtually all Gaelic speakers are now bilingual, their presence increases the likelihood that any given situation will be

²⁶ Since writing this, two novels have appeared which happily go against this tendency: *Gymnippers Diciadain* by Màrtainn Mac an T-Saoir, 2005, and *Shrapnel* by Tormod Caimbeul, 2006

English rather than Gaelic-medium. However, even in conversations between fully fluent Gaelic speakers, code switching and borrowing is currently the rule rather than the exception.

There is a wide range of print and media (i.e. television and radio) registers available in the language, but the tendency is for there to be generally few instances, especially in printed material, of each specific variety. This gives the overall impression of breadth without depth in register range. Although a greater number of persons returned in the 2001 census as being literate than conversant in Gaelic, very few find opportunities to write or read the language on a regular basis. This small readership is a contributing factor to the dearth of prose in all forms in the language, especially long fiction and academic writing. In fact, a substantial proportion of Gaelic writing, in particular children's books, news reportage, and government documents, are translated from English publications. As a result, cross-linguistic effects are readily observed contributing to, on some level, the growing departure in the language from Gaelic-based idiom towards a higher instance of calquing and decay of native morphosyntax.

4 Methodology

4.1 Data Collection and Composition

Leech (1991) states that the main difference between a 'corpus' and an 'archive' is that the former is a collection of text for some *purpose* while the latter is simply collection for its own sake. The primary interest of the current study is to investigate linguistic register differentiation in Scottish Gaelic. A corpus constructed for this purpose ought to be representative of the kinds of variation found in the language at large, a sample of a greater population. Due to time and resource constraints it was possible to sample a core group of only four spoken and four written registers. However, these were carefully selected to be as representative as possible of the broad band of linguistic and contextual variation available in Scottish Gaelic.

The initial set of registers is based upon the results of previous research (especially Biber 1988; 1995; Chafe 1982 *inter alia*; Ferguson 1983; Greenbaum *et al.* 1995a *inter alia*; Jang 1998; Miller and Weinert 1998; Romaine 1994) as well as a consideration of which registers had specific cultural importance to Scottish Gaelic.

In accordance with the relative ease or difficulty of obtaining the different types of text, a final set was chosen consisting of:

<i>Spoken Sub-corpus</i>	<i>Written Sub-corpus</i>
1. Conversation	5. Academic Prose
2. Radio Interview	6. Fiction
3. Sports Broadcast	7. Popular Writing
4. Traditional Narrative	8. Radio News Scripts

There was only a limited amount of time to collect and tag the corpus. It was crucial for it to be of a manageable size, but one that would not compromise its validity. Biber (1990) investigates a number of methodological issues regarding corpus-based studies on language variation, in particular the impact of text length, category size (i.e. number of texts), and corpus size on validity. As his independent variables, he takes a small number of linguistic features that have been commonly used in language variation research. He then assesses their distribution throughout sub-samples (50% and 25%) of the Lancaster-Oslo-Bergen Corpus of British English and the London-Lund Corpus, in comparison to their distribution in the full versions of these corpora. His results indicate that categories (i.e. register samples) consisting of ten ($n=10$) 1000 word texts are sufficiently large to investigate these kinds of variables. In other words, despite the fact that these sub-corpora were small samples of larger pools of data, they were still representative of the features investigated. According to Biber, size is important²⁷, but more crucial is the corpus being representative of the language varieties that it purports to reflect.

In line with Biber's findings and deeming it a manageable goal, the average length of individual texts in the present research was set at 1000 words. These grouped into 10-text subcorpora to result in an overall corpus of 81,677 words. This is comparable to the Leverhulme Corpus (42 texts of 2000 words, 84k total) used by Greenbaum and Nelson (1995) to investigate clause relationships in English. It is many times larger than Nivet's (1989) study of the syntax and discourse of conversational French, based upon 600 'sentences' per speaker. If spontaneous speech is defined as non-scripted spoken language, the current data is twice that used in Miller and Weinert's (1998) syntactically-coded corpora of Scottish English. On the other hand, it is smaller than Macaulay's (1991) study of Scottish English (~110k words). All of these studies have been different in focus, and each has produced interesting results. What they have in common is the manual coding of

²⁷ It is particularly important for lexicological studies.

syntactic and discourse-related features, without the use of computerised taggers. Although these studies employ much smaller corpora than most multidimensional approaches, they achieve a much more in-depth analysis of the phenomena to which they attend.

Originally, the intention of the present research was to do a multi-dimensional study of Scottish Gaelic, based upon Biber's methodology. After consideration, it was determined that the time it would take to write a programme to carry out the grammatical tagging would be better spent doing it by hand. So far, multi-dimensional studies have relied on the frequency counts of fairly simple and prevalent morphosyntactic features (cf. Miller and Weinert 1998: 14). More complex or subtle features, such as those implicated in information structure, are generally ignored because they are not easily detected by computerised search routines. In deciding to carry out a manual analysis, a smaller body of data is covered than an automatic tagger would have permitted. However, as will be clear, many of the most interesting results of this study would have been precluded in an automated approach. In addition, it would have been difficult if not impossible to write a programme that would have been able to make many of the distinctions that are needed.

4.1.1 *Origin and Selection of Texts*

The data is from a number of different sources. The table below and following discussion summarise the overall composition of the corpus. For information on specific texts, please refer to Appendix 2.

Table 4: Organisation and Make-up of the Scottish Gaelic Corpus

<i>Spoken Sub-corpus</i>	<i>No. of Texts</i>	<i>No. of Words</i>	<i>Sources</i>
1. Conversation	10	11,501	b
2. Radio Interview	8	10,034	b
3. Sports Broadcast	10	10,278	b
4. Traditional Narrative	10	10,221	a b c
<i>Written Sub-corpus</i>			
5. Academic Prose	9	10,089	a c
6. Fiction	8	9,630	c
7. Popular Writing	11	9,690	a c
8. Radio News Scripts	10	10,234	a
Totals	76	81,677	

Sources: a) electronic sources (i.e. web pages, computer files); b) transcriptions of recorded material; c) scanned documents converted to texts using optical character recognition (OCR) software

Conversation Texts

These were recorded on minidisc or digital audio tape using a small stereo microphone. They were then dubbed onto audio tape for transcription by the author and Kirsty Mac Donald, a native speaker from North Uist. The participants were mostly middle aged, except for three children in two related texts, and all had acquired their Gaelic through intergenerational transmission. The recordings took place in domestic settings, between close familiars and family members, mainly over tea or during supper. The participants were aware that they were being recorded but, as indicated by the affect level of some recordings and the private nature of many topics (such as sensitive gossip), they did not seem inhibited by this fact. For some recordings WL was present but in most the recorder was left in the care of the participants.

Radio Interviews

These were all broadcast on *Radio nan Gaidheal (RnG)* between 1980 and 2000 and were taped either by WL or by BBC employees. Five of the eight programmes were from the morning show 'Coinneach MacIomhair'. In this programme, Mr. MacIomhair acts as an interviewer-mediator to discussions on a wide range of current-interest topics, with both phone-in and studio-based participants. The other three programmes had similar formats. All texts were transcribed by WL and KL.

Sports Broadcasts

These consisted of on-line coverage of two team Scotland football games (*vs.* Australia and Yugoslavia). There were two presenters, one covering action on-field and the other summing up between plays. These were broadcast on RnG and recorded by BBC personnel. The transcriptions were made by WL and KL.

Traditional Narrative

All of the traditional narrative recordings were from the archives of the School of Scottish Studies, collected from Uist-born informants²⁸ during the late 1960's and early 1970's. An effort was made to include a wide range of different genres such as international fairy tales, Gaelic legends, hero tales, and local history. Many of the narrators were elderly when recorded, thus age factors are of consequence to the interpretation of these texts. Some of these were transcribed by WL, others were published in books (see Appendix 1), while others were in the form of text files, having been transcribed by Peggy Clement at the School of Scottish Studies. All texts which were not transcribed by WL were checked for accuracy using copies of the original recordings.

Academic Prose

Academic prose was defined as expository writing by teachers and researchers involved in tertiary education about topics most characteristically appearing in pedagogical contexts: poetic analysis, biology, linguistics, Scottish history, and so on. Most of the texts have been published in books or as chapters in compilations. One text is the introduction from an unpublished PhD thesis. All texts, save the latter, were scanned into electronic form and converted to text files using an OCR package.

Fiction

These were mostly short stories, except two, which were the first chapters of longer novellas. All were published. Several different authors were included, writing on a wide range of topics (see Appendix 2). These were all scanned and converted to text files using OCR software.

Popular Writing

Popular writing was defined as in Jang (1998: 45): "short expository essays written for the general public". All of the texts included were published as general interest articles in newspapers such as *The Scotsman*, and *An Gaidheal Ùr*. They concern a wide range of topics from religion, to reviews, to reminiscence. Some were retrieved from web pages while others were scanned into the computer and converted to text.

²⁸ Largely, in the case of those I used, by Angus John MacDonald. Currently a Gaelic teacher in Portree High School, AJ was, some say, the best collector the School of Scottish Studies ever had. It is hoped that some day AJ publishes the monumental work that he did for his unfinished PhD thesis and receives the credit that he is long due.

News Scripts

The news texts are from RnG workers during the late 1990's. These were sent to WL in the form of e-mail file attachments from the BBC. Pieces from most of the news staff that were working for the BBC Gaelic Department at the time are included and they are fairly equally represented in the subcorpus. Spelling errors were corrected by WL before inclusion in the corpus.

4.1.2 Transcription Practices

With every transcription, one must consider which features of language are to be represented, how it is to be used, and who are its intended users (Crowdy 1995). For this study, only a simple orthographic transcription was needed and anything beyond this would have been both superfluous and overly time-intensive. Spoken texts were transcribed without punctuation (except question marks to indicate interrogative intonation) and no attempt was made to indicate pause length. Filled pauses (*um*, *hà*, *ehm*, etc.) were not included, but otherwise, the aim was to produce a faithful reproduction of the lexical content of the recordings. Orthographic practices were broadly consistent with the recommendations of the Scottish Examinations Board (SEB).

Speaker turns were indicated in the transcriptions by [1], [2],²⁹ ... etc. Where significant cross-talk, low recording level, or other factors conspired to prevent accurate transcription of content, a marker '[?]' was used. Code switching was always included in the transcribed text and tagged as such.

4.1.3 Situational Description of the Registers with Examples of Texts

Registers are language varieties that correlate with certain contextual parameters. A prerequisite to understanding the distributions of linguistic features across the different types of registers is to understand the situational factors that help condition this variation. Biber (1994; cf. 1988: 29–33) proposes a number of different components of the speech situation based upon prior research on the ethnology of communication acts. This framework is employed in the following discussion, which describes and makes some comparisons between the contextual aspects of each register included in the study.

4.1.3.1 Conversation

The number of participants present in the conversations ranged from two to eight. In each recording, all were close friends or family members and, in most cases, were roughly the same age (45–60 years). As mentioned above, two texts were from a family with young children (9–15 yrs.). All participants were of a similar economic background, broadly characterised as middle class. The occupations of the adults were mainly crofting, education (primary and secondary) and banking. Adult participants varied somewhat in the educational level they had attained, from secondary certification to postgraduate degrees. Except in the texts including the children, all participants would have had roughly equal status. This is a crucial feature distinguishing natural conversation from nearly all forms of public or radio conversation. Their shared world knowledge is evident in the texts, with frequent references to persons, places and occurrences that are part of their common experience. The topics covered in the texts include the breadth expected from spontaneous conversation between familiars: inquiring about each others' days, gossiping about neighbours, discussing politics, requesting food or drink, and so on. Some were framed by a superordinate activity, such as sitting down for a meal, yet in others the activity was limited to conversation. In all cases, the recordings were made in a domestic setting. One can suppose that the purposes of intimate conversation have to do with group consolidation, entertainment, information exchange, and obtaining assistance. A

²⁹ These are presented here, in the transcriptions provided, as [a], [b] etc. to avoid yet another number set.

cynical commentator might suggest that the purpose of conversations that are being taped is to gratify the researcher, but none of the recordings had pervasive or long-lasting signs of artificiality. Where initial signs of self-consciousness were present, they dissipated within minutes. For some recordings, the verbal consent of those participating was gathered before the taping ensued, while for others this was done *ex post facto*.

Example of conversation text:

- <a> *nach do rinn iad an aon rud oirne an council a bhith 'gealltainn nach cuireadh iad charge air son an sewer a bha siud nuair a bha iad ga thogail 's an ath rud a-nuas a thàinig NOSWA a-staigh 's bhuail iad air na rates*
- *na sewer rates*
- <a> *co-dhiubh tha thu 'join-adh gus nach eil*
- <e> *dè chuir iad air*
- <a> *och tha e brùideil co-dhiubh*
- *nach eil e còrr 's ceud nòt*
- <a> *tha e a' sìor-dhol suas*
- <d> *rud a th' agamsa chan eil còir aca charge-adh air co-dhiubh chan eil ann ach outside tank tha còir a'd air aon tank fhaighinn*
- <a> *'s e tha siud ach an trough ge-tà*
- <d> *ach chan eil trough idir againne*
- *uill chan eil còir agads' a phàigheadh idir ann ma-thà cha leig thusa leas a phàigheadh ma-tha feumaidh tu a chantail riutha*
- <a> *'son am beagan a thathar ag use-agadh a dh'uisge*
- <d> *faodaidh an hose a bhith an àite sam bith*
- <a> *uill faodaidh tu aona tap a bhith agad co-dhiubh ma-thà*

Translation:

- <a> didn't they do the same thing to us the council promising that they wouldn't levy a charge for that sewer when they were building it and the next thing up comes NOSWAS and they slam on the rates
- the sewer rates
- <a> whether you join them or not
- <e> what did they levy
- <a> oh it's brutal anyway
- isn't it more than a hundred pounds
- <a> it's constantly going up
- <d> the thing I've got they don't have a right to charge for it anyway it's only an outside tank you've got a right to get one tank
- <a> that's the trough though
- <d> but we don't even have a trough
- well you shouldn't be paying it at all then you don't have to pay then you have to tell them
- <a> for the little that people use of water
- <d> the hose can be anywhere
- <a> well you are able to have one tap anyway then

4.1.3.2 Radio Interview

In the radio interviews included in the study, there are addressees, generally one addresser at a time, and an audience removed physically and, in the case of re-broadcasts, temporally from the context of language production. Turns are negotiated by the interviewer, who has more status than the interviewees and is responsible for controlling and framing the discussion so that it is interesting, provocative, enjoyable, and

informative for the audience. The interviewees are occasionally able to steer discussion themselves, although their efforts are in deference to the interviewer and, thereby, transient. In some cases, the setting may be solely a studio at the radio station, but it is usually multifarious with several participants included in the discussion by phone. Because of this, there is often a combination of face-to-face participants and those who, because they are on the phone, are unable to either transmit or receive non-verbal communication. The interview is usually moderately structured (i.e. some questions are prepared beforehand) and the interviewees often know which questions they will be asked to begin with, allowing them time for preparation as well. As a result, there is the potential for more literate language forms to enter unseen into the dialogue. The attitudinal stance of the participants to the text can vary depending on the topic and the level of emotion that it invokes in them. The interviewer is expected to remain neutral although s/he may, through leading questions and other rhetorical devices, influence the discussion for effect. The purpose of this kind of discourse can vary, but tends towards the entertainment and edification of the audience. The flow of conversation is always self-consciously geared towards this orientation. The level of discourse is usually of a popular nature, readily appreciated by the listeners, but with a wide range of specific subjects represented.

Example of radio interview text (Coinneach Mòr's show):

- <a> uell còmhla rium tha a's an studio ann a-seo an-dràsta tha Murchadh Dòmhnallach tha cù
snog aige 'na chois an-sin cù agus h-abair g'eil e dòigheil a Mhurchaidh ach tha sgeul co-
cheangailte ris a' choin [sic] nach eil
- tha 's ann à Uibhist a thàinig an cù agus ' se rescue dog a bh' ann agus cha robh iad cha
robh iad a' toirt dha biadh no càil
- <a> dè thachair dè a' suidheach: staid a's a robh e nuair a fhuair thusa e
- bha cha robh cha mhòr gu robh bit feòil air
- <a> seadh seadh
- bha e dìreach bha an staid a bh' ann uabhasach 's bha a chalg a' falbh is a h-uile càil
- <a> agus dè mu dheidh: dè mu dheidhinn man a bha e le daoine 's mar sin
- bha e gu math feagalach nuair a thàinig e a-seo ach an-diugh thà tha e tòrr nas fheàrr na
bha e

Translation:

- <a> well along with me in the studio here right now is Murdo MacDonald he has a nice
dog with him at his heel there a dog and isn't it nice Murdo but there is a story
attached to the dog isn't there
- yes it's from Uist that the dog came and it was a rescue dog and they weren't giving
him food or anything
- <a> what happened what state was he in when you found him
- there was almost no flesh on him
- <a> right right
- he was just the state he was in was horrible and his fur was gone and everything
- <a> and what about how was he with people an so on
- he was very fearful when he came here but today he is much better than he was

4.1.3.3 Sports Broadcast

In many ways, this register is similar to radio interview: there is an addresser, an addressee, and an audience who is removed, physically and possibly temporally, from the production circumstances. The addresser will always be either the play-by-play presenter, who is responsible for action on field, or the inter-play commentator, who sums up, opines, and draws associations for the benefit of the listening audience. Turn taking is dictated by on-field action. Insofar as action description is deemed more important than

commentary, the play-by-play presenter has more status than the inter-play commentator. This can be witnessed primarily in interruption behaviour, with the former taking the floor if action resumes even before the latter is finished speaking. The immediate setting of language production is a small box or studio within view of the playing arena, equipped for radio broadcast or recording. There are extreme but different production constraints on the addressers: the play-by-play presenter must describe each action on field as it occurs with little time for elaboration. Thus he has a relatively superficial task, but one requiring quick and concise words. The inter-play commentator has a more cognitively demanding task, and one that must occur in the gaps between on-field action. The addressers' stance towards the text can range from neutral to emotionally involved (particularly at times of scoring and penalties). The addressers sympathise with and revel along with the audience at peak times of play. The main text is based on fact – action that is occurring in real-time. Similar to radio interview, the purpose is primarily to entertain and inform. Unlike radio interview, the level of discussion is specialised and it involves two main subjects: football in general and the particular football game in question.

Although spontaneous football reportage is not a translation-based register, like news language there has been a conscious attempt towards establishing a Gaelic-based lexicon that would be appropriate for the radio. The following short glossary is from Angus Macdonald at *Radio nan Gàidheal*, in an e-mail to the author dated 20 March 1999. He said that the sports announcers “generally made an attempt to use words in existence in a different context, plus a few [new] phrases ...”:

<i>bròg</i>	a good belt of the ball [Lit. ‘a shoe’]
<i>buille bhon oisean</i>	corner kick
<i>buille peanasach</i> [sic]	penalty kick
<i>cluicheadair adhart</i>	attacker
<i>cluicheadair dìon</i> [sic]	defender
<i>cluicheadair meadhan</i> [sic]	midfielder
<i>dhan a' mhòintich e</i>	over the stand clearance [Lit. ‘to the moor it’]
<i>dheh</i>	off-side [Lit. ‘off’]
<i>fear a' loidhne</i> [sic]	linesman [Lit. ‘man of the line’]
<i>fear-glèidhidh</i>	goal-keeper
<i>rèitire</i>	referee
<i>ro-fhada suas</i> [sic]	off-side [Lit. ‘too far up’]
<i>sadadh a-steach</i>	throw-in
<i>tadhal</i>	goal
<i>tilgeadh a-steach</i>	throw-in
<i>uinleag</i>	‘ungentlemanly elbow’

Example of sports broadcast text (Scotland *vs.* Australia):

<a> Aitken gu Macavenie Macavenie air ais gu Malpas tha i dìreach [?] a-staigh air leth na h-Astràilianaich [sic] gu Dalgleish a's a' bhucais [sic] agus fhuair Creeno a chas uice a-rithist agus chuir a-mach i air taobh thall na pàirce agus 's e sadadh a-steach a bhios ann an siud a-rithist do dh'Alba 'se Aitken fhèin a tha 'dol dha gabhail tha e ga feuchainn gu Dalgleish ruith Dalgleish a-null Dalgleish a' chiad turas a-null a-null dhan a' bhucais agus chaill Greedy i ach bha Macavenie bha e dìreach slat no dhà air dheireadh chaidh siud airson buille bhon oisean a Mhurchaidh

 ma chumas seo a' dol Aonghais chan eil againn ach dìreach feum 'son tadhail gheibh sinn gheibh sinn fear gheibh sinn fear a dh'aithghearr tha mi a' smaoineachdainn tha am ball a' dol a dh'àite sam bith a-nis aig ga shadail a-mach no ga bhualadh a-mach a dh'àite sam bith air cùl na pùisd no a-mach dhan taobh eile na pàirce

<a> Strachan leatha gu Cooper...

Translation:

- <a> Aitken to Macavenie Macavenie back to Malpas it's exactly [?] in on Australia's side to Dalgleish in the box and Creno got his foot to it [the ball] again and put it out on the far side of the park and it's a throw in that there'll be there again to Scotland it's Aitken himself that is going to take it and he is trying it to Dalgleish Dalgleish ran over Dalgleish the first time over over to the box and Greedy lost it but Macavenie he is just one or two yards away behind that went for a corner kick Murdo
- if this keeps going Angus we only need a goal we'll get we'll get one we'll get one soon I am thinking the ball is going anywhere at all now thrown out or kicked out to anywhere at all behind the post or out to the other side of the park
- <a> Strachan with her [the ball] to Cooper...

4.1.3.4 Traditional Narrative

The tapes used for this project all have a single addresser, and one or more addressees (one of whom could be considered the tape recorder itself). They were recorded usually in the homes of a raconteur/storyteller by an interviewer, a worker from the School of Scottish Studies trained in interview and field-recording techniques. The relationship between these two participants can be characterised as mutual respect and, in some cases, friendship, however the interview has a dominant influence on the progression of discussion topics following or preceding each piece of narrative. The storytellers would have been cognisant of the fact that they were being recorded for study and archival purposes; that there would eventually be an 'audience' removed in time and space. While the language of the storyteller is produced on-line, there is ample anecdotal and textual evidence to conclude that much of the language in the stories was crystallised to some extent. This occurred through numerous hearings and retellings and, in some cases, the stories and their language can be traced back to literary manuscripts that were in circulation up until the 18th and, rarely, the early 19th century (see Bruford 1969; Dòmhnallach 1989). Many of the storytellers became excited when narrating, to the extent that one could say they were emotionally involved in the subject matter. Some professed the conviction that the events occurring in the stories, no matter how apparently chimerical or fabulous to modern sensibilities, actually took place. By and large, recognised narrators of traditional stories value their material greatly; if they did not, they would not have been likely to have listened to and remembered the material from other narrators. As Akinnaso mentions (1985, 346) the conservative form of such 'ritualised' texts is partly due to their repetition over a long period of time, but also a factor of the reverence reserved for them in their culture. Although it is difficult to pin down the purpose of such a wide range of discourse, traditional narrative certainly functions to entertain and transfer information. The level of discussion (e.g. specialised/general/popular) varies according to subject matter, which itself can vary widely. However, a great deal of shared knowledge usually exists between the storyteller and the addressee(s), characteristic of their common social background or, in the case of learned interviewers, their academic study of Gaelic folklore.

Example of traditional narrative (from Moireasdan 1977: 53):

Tha sgeulachd agam ri innse dhuibh a nochd agus tha mi 'smaoineachadh gun creid sibh gun gabhadh e tachairt, thoradh àite sònraichte bha na bhonn stéidh air an deach a togail an toiseach, tha e ri fhaicinn gun a' latha 'n diugh: ged nach fhaic sinn leis an t-sùil e no le glain'-amhaire tha e ann – creag mhór àrd dà cheud mìle muigh 'sa' Chuan an Iar o thaobh an iar Uibhist ris an cainte o chionn fada sa Ghàidhlig 'Ròcabarraigh' ach sa chànan Bheurla 'n diugh Rockall, 'ciallachadh gur e stall' a th' ann uileag 's nach eil feur idir orra. Mar a tha 'n t-seann sgeulachd a' dol, bidh iomadh linn on a ghabh i àite ach tha Icho ùr dhuinne bhith ga h-innse 'n diugh, no bhith 'g éisdeachd rithe o neach eile ga h-innse, 's nach

eil sinn a' faighinn atharrachadh sam bith innte on a chuala sinn an toiseach i. 'S e "Bodach Ròcabarraigh" is ainm dhan sgeulachd.

Translation:

I have a story to tell you tonight and I am thinking that you will believe that it could have happened, because of a particular place that was the basis on which it was made, it can still be seen to this day: although we can't see it with the eye or with a telescope it exists – a big high rock two hundred miles out in the Atlantic Ocean from the west side of Uist which was called since a long time ago in the Gaelic 'Ròcabarraigh' but in the English language today 'Rockall', meaning that it is all a sea rock and that there isn't grass at all on it. As the old story went, it will be many generations since it took place but it is so new to us telling it today, or to be listening to it from another person telling it, that we don't find any change in it from when we first heard of it. "The Old Man of Rockall" is the name of the story.

4.1.3.5 Formal Prose

Written prose is generally characterised by having usually one addresser and an unlimited number of addressees. Because the writer has control over the text production, s/he can be said to have more status than any addressee. Formal prose, in contrast to most other kinds of writing, is associated with the sharing of specialised knowledge; addressers write in a way that assumes addressees understand the jargon and conceptual territory of the discipline in question. Place and time are not shared by the participants and the record is permanently written and usually published in the form of books, chapters in books, or articles in academic journals. The use of electronic media (webpages, downloads) is growing as a venue for the register, particularly for so-called 'working papers', essentially pre-publication versions. Authors are able to revise as much as they wish. Published formal prose has usually been rigorously edited and commented upon by peers in the writer's own discipline to the extent that, in some ways, the form becomes the product of several individuals while the message is the product of one or few. Both addresser and addressee generally view the text as important and valuable. The addresser's stance toward the text tends to be emotionally removed but s/he will be strongly convinced of its veracity. Formal prose can be persuasive to varying degrees but is nearly always oriented towards the transfer and exposition of information. The level of discussion is specialised but the specific subject concerned varies according to discipline.

Example of academic prose (from Dòmhnallach 1989: 185):

A nis, ann an Eirinn agus ann an Albainn faodaidh sinn a ràdh gun do mhair an saoghal, ann an iomadach dòigh, 'meadhon-aoiseil' mòran nas fhaide na rinn e ann an cuid de àiteachan eile. Bha seo mar thoradh air tachartasan eachdrachail agus poilitigeach nach ruig mi a leas a rannsachadh an dràsda. A bharrachd air a seo, faodaidh sinn a ràdh gun do dh'fhan Gàidheil Eirinn agus Alba nan aona nàisean, a thaobh cultuir, suas gu an t-seachdamh ceud deug, agus eadhon gu an latha an diugh fhéin ann an iomadach dòigh; agus tha seo gu h-àraidh fìor a thaobh an cuid sgeulachdan. Gu dearbha, fhad agus a mhair an ceangal dlùth eadar an dà dhùthaich, suas gu an t-seachdamh ceud deug, bha cuid de Ghàidheil Alba a' coimhead air Eirinn mar an 'Dùthaich Mhàthaireil', agus bha seo gu h-àraidh fìor a thaobh nan seanchaidhean agus nam filidhean ionnsaichte a bha, nan cuid sgrìobhaidh, a' cleachdadh an aon seòrsa de Ghàidhlig chlasaigeach agus sheann-fhasanta agus a bha na seanchaidhean Eireannach a' cleachdadh.

Translation:

Now, in Ireland and in Scotland we may say that the world lasted, in many ways, 'middle-aged' much longer than it did in some other places. This was a result of happenings historical and political that I need not examine right now. On top of this, we

may say that the Irish and Scottish Gaels stayed as one nation, in terms of culture, up to the seventeenth century, and even until today in many ways; and this is especially true concerning stories. Indeed, as long as the close connection between the two countries lasted, up to the seventeenth century, some Scottish Gaels looked at Ireland as the 'Mother Country' and this was especially true with the tradition bearers and educated poets that were, in their writings, using the same type of classical, old-fashioned Gaelic that the Irish tradition bearers were using.

4.1.3.6 Fiction

Like formal prose, fiction most often originates with one addresser, is highly edited, and usually published in books, book compilations, or journals. Unlike formal prose, shared specialist knowledge is not normally necessary between the addresser and addressee(s). The writer may not be perceived as being more socially powerful than the reader in the context of the reading activity, but in a wider sense, writers are often thought to be 'special', often in proportion to the extent of their fame. Unless the work is biographical, the addresser does not usually share personal knowledge, at least not in a transparent way. The addresser may personally evaluate the text as important, valuable, personal, or even beautiful. Addressees may or may not have strong feelings to the text, but fiction perceived as powerful may evoke, by proxy, the whole gamut of emotion and cogitation available to human experience. The purpose of fiction centres around entertainment as well as expressing and instilling imagination, but may be coloured by degrees of persuasion, self-revelation, or transfer of information. Although genre fiction such as sci-fi can have a very specific level of discourse, most fiction is accessible to a popular audience.

Example of fiction text (from MacAonghais 1979: 1)

"Bìdh mi air ais mun dorchnaich i," thuirt e mun do dh'fhàg e 'sa mhadainn. Thàinig i nuas cuide ris gu ceann an rathaid.

"Ma bhitheas, ' se chiad uair dhuìt e," ars ise. " 'S neònach leam-sa mura téid thu shuiridhe dh'àiteigin air do thilleadh, mur do dh'atharraich thu. Ach mura nochd thu roimh mheadhon-oidhche, glasaidd mi 'n dorus ort."

Rinn e gàire. ' Se dh'fhaodadh, is fhios aige nach dèanadh i dad dhe leithid. 'S iomadh uair, nuair a bha e dol do sgoil an Rubha agus e loidseadh aice, a bhiodh e muigh gu uair 'sa mhadainn; 's cha robh an dorus glaiste riamh.

"Feuch a neisd, 'ille, nach beir an làn ort. Ma chailleas tu an fhadhail, cuir fòn thugam a Post Office an eilein. Bi cinnteach, a neisd, ma chailleas tu do shuim nach toir thu oidhirp air an fhadhail, mus bàthar thu. B 'fheàrr dhuìt cus fuireach thall gu madainn na thu fhéin a chur an cunnart a' greasad dhachaigh. Cuimhnich a neisd!"

Translation:

"I'll be back before it gets dark," he said before he left in the morning. She came down with him to the end of the road.

"If you are, it'll be the first time for you," she said. "I'd find it strange if you didn't go off courting somewhere after you return, unless you've changed. But if you don't appear before midnight, I'll lock the door on you."

He laughed. So he would, and him knowing that she would never do the like. Many's the time, when he was going to the Point school and lodging with her, that he would be out until one in the morning, and the door was never locked.

"Mind now lad that the tide doesn't catch you. If you lose the ford, give me a call at the island post office. Make sure now, if you lose track of time, that you don't attempt the ford, in case you drown. It would be better for you to spend too much time over there until the morning than to put yourself in danger hurrying home. Remember now!"

4.1.3.7 Popular Writing

There are several ways in which popular writing diverges from fiction and academic prose. First, the register is intended for a very general audience: popular writing is found mainly in newspapers, which are available in public shops and thus must cater to a wide readership. Second, newspapers are designed for transience; they are read in the day or two after they are printed and then discarded. This fact is reflected in the length, subject matter, and level of editorial polish found in this kind of writing. Third, popular writing is often embedded within a larger surrounding register: news reportage. Fourth, because newspapers are published regularly, often daily, the addresser and addressee of popular prose are temporally closer than non-newspaper writing. The addresser of this register may value it, but not generally to the same extent as more permanent discourse types. As s/he may write weekly or even daily, any attitudinal stance toward the text is likely to be conditioned by its evanescence. The purpose of popular writing can vary widely: a single piece can veer between persuasion, exchange of information, entertainment and self-revelation. Almost always, the level of discourse will be 'popular' with a specific subject determined in part by the forum's characteristics, current affairs, and the recent personal experiences of the writer.

Example of popular writing text (from NicDhòmhnaill 1999):

Bha Celtic Connections 1999 a' dol air adhart na bu tràithe air a' mhìos seo ann an Talla Chonsairt Rìoghail Ghlaschu.

B' e seo an t-siathamh bliadhna dhen taisbeanadh seo de shàr luchd-ciùil agus ealain nan dùthchannan Ceilteach gu léir, le deagh mheasgachadh de cheòl, de sheinn, de dhannta, de sheanchas, de dheasbairteachd agus eile fad naoi latha deug, le còig consairtean gach latha. Faodaidh sinne, mar Ghaidhlig, beachdachadh air ciamar a sheas Gàidhlig ghrinn na h-Albann agus a cultar an coimeas ri cultaran nan dùthchannan eile.

Mar eisimpleir, air dhomh sùil aithghearr a thoirt thro leabhraan Celtic Connections, bha e soirbh a dhèanamh a-mach nach robh uidhir sin dhe ar cultar fhèin air a riochdachadh an taca ri cuid de na dùthchannan eile.

Translation:

Celtic Connections 1999 was going forward earlier this month in the Glasgow Royal Concert Hall.

This was the seventh year of this exhibition of excellent musicians and artists of all the Celtic countries, with a good mixture of music, singing, dancing, oral traditions, debate and other items for nineteen days, with five concerts every day. We may, as Gaels, opine on how the lovely Gaelic of Scotland stood, and her culture, in comparison to the cultures of other countries.

As an example, after I had a quick look through the booklet of Celtic Connections, it was easy to make out that there was an unequal proportion of our own culture represented compared to certain other nations.

4.1.3.8 Radio News Scripts

There are several ways in which radio news scripts diverge from the other registers used in this study. First, although they are produced in written form, they are transmitted orally. In essence, there are two texts: a script which is never seen by the public and the on-line speech which, unless it is recorded, has no permanence. Second, the writer works under time-pressured production circumstances: each in-coming news item (off the wire) must be translated from English to Gaelic with very little time for revision, or editorial control.³⁰ Third, the information contained in the news scripts, like other forms of journalism, is of transient importance: news must be 'new' and it is valued mainly for its

³⁰ Jang (1998) reports exactly the same circumstances for Taiwanese news scripts, which are translated from Mandarin Chinese and exhibit a heavy influence from the source language.

accuracy and promptitude. Comprehension circumstances are on-line: the addressee is unable to stop the addresser for clarification. Because of this, there is an onus on the addresser to present the information in a manner allowing easy perception and comprehension, a goal that is sometimes obviated by production circumstances. The addresser may consider the text to be important or of value but, as there is little room for creativity, s/he is not likely to evaluate it in an extended aesthetic sense. The addresser's attitudinal stance towards the text is likely to be emotionally removed, in balance with the purpose of the register – to transfer information – and its overriding concern with factual *versus* imaginative material. The level of discourse is usually general and the specific subjects covered will be a function of what is considered newsworthy by the agencies supplying the 'wires'. Finally, the news script register shows a steady forward motion, with little returning to previous topics or referents. In this way, it diverges from more discursive registers.

Example of radio news script (UNEDITED):

Tha na polais a tha rannsachadh murt Thomas Marshal, ag radh gu bheil uad [sic] airson faighinn a mach an robh caraid diomhair aig a bhalach, gun fhiosda dha theughlach [sic]. Chaidh an corp aig Tomas, a bha dusan bliadhna dh'aois, a lorg faisg air baile Thetford ann an Sasuinn Di-Huoinne 'sa chaidh. Chaidh e air dhith bho dhachaidh ann am Habbisburgh air an latha roimhne sin.

Theid coinneumh diomhair a chumail ann am Port Righ le Comhairle na Gàidhealtachd an diugh, son deasbud a dheanamh mu seann phort adhair An Ath Leathainn. Bha'n comhairle [sic] air diuileadh roimhne seo cead a thoirt do chompanaidh a Sasuinn am port adhair a chleachdadh son a bhi a ruith [sic] cuairtean helicopteran dhan a Chuilfhion [sic]. Tha'n comhairle [sic] air diuileadh [sic] dad a mu de a th'aca san amharc son a phort adhair.

Translation:

The police are investigating the murder of Thomas Marshall, saying that they want to find out if the boy had a secret friend, unknown to his family. The body of Thomas, who was twelve years old, was found near Thetford in England last Friday. He went missing from his home in Habbisburgh the day before that.

A closed meeting will take held in Portree by the Highland Council today, in order to discuss the old Broadford airport. The Council has refused before this to give permission to a company from England to use the airport to run helicopter trips to the Cullins. The Council has refused to say what they have in plan for the airport.

4.2 Selection of Linguistic Features

The linguistic properties surveyed in this research originated with a number of sources. Because of the dearth of work in relevant areas on the Gaelic language (see Chapter 1), possible points of investigation were mainly located in work on other, more surveyed languages (especially English). Amongst the most helpful of these were the studies of Biber (especially 1988 and 1995) and others adopting his multidimensional framework (Besnier 1988; Kim and Biber 1994; Jang 1998) as well as the work of Chafe (1982 *inter alia*), Macaulay (1991), Lambrecht (1994), Greenbaum and Nelson (1995a *inter alia*), and Miller and Weinert (1998). An initial read through of a wide range of different ScG texts also provided some possibilities. However, many of the features that were eventually settled upon only arose out of the actual tagging process. As a result, there was considerable revision of the tag set in the early stages both to incorporate what were perceived as important trends in the data set, and to dismiss features that were proving to be less salient in the greater context of the study. In general, the aim was to include as many features as possible in the initial mark-up, for descriptive comprehensiveness and to have the widest possible net for catching potentially important distinctions between the texts.

The following tables and discussion detail the features that were settled upon and used as the basis for the quantitative analyses. The full tag set contains more features than are included here (see Appendix 3 for an exhaustive list of all the tags used). Some of these will be discussed in later sections but were excluded from the quantitative study due to low overall frequencies. Many of the features below were scored by combining various discrete tags. Relevant sections in the descriptive grammar have been provided for further information. In Chapters 5–7: information is provided on the distinctions of each feature and their importance to the study, with references to past research and examples. A list of the most important features in the study, along with their contextual/functional correlates, is provided in the introduction preceding the results chapters (p. 71).

Table 5: Corpus tags used for the quantitative analysis

A. Information Structure and Discourse

Feature Name	Example(s)	Grammar Section
1. Clefts: <i>'s</i> -type	<i>'s ann an-shin a bha e</i> it was there <i>that he was</i>	3.4.2.2
2. Clefts: <i>se</i> -type	<i>'s e fhèin a chaidh ann</i> it's he himself <i>who went there</i>	3.4.2.2
3. Clefts: deictic	<i>sin an duine a bha ann</i> that's the man who was there	3.4.2.2
4. Detachment: left-detached elements	<i>Seumas dh'fhalbh e mar-thà</i> James <i>he left already</i>	3.4.2.2
5. Detachment: right-detached elements	<i>tha e grànda an taigh ud</i> <i>it is ugly</i> that house	3.4.2.2
6. Placement: fronted adverbials and peripheral elements	<i>a-mach a ghabh iad</i> out <i>they went</i>	N/A
7. Placement: postposed subordinate clause	<i>cuir an cù a-mach nuair a dh'fhalbhas tu</i> <i>put the dog out</i> when you leave	3.3.2.2
8. Placement: preposed subordinate clause	<i>nuair a dh'fhalbhas tu cuir an cù a-mach</i> when you leave <i>put the dog out</i>	3.3.2.2

B. Clause Types

9. Coordinate clause	<i>... agus tha sibhse a' fuireach</i> <i>... and you are staying</i>	3.3.1
10. Cosubordinate clause with <i>agus</i> 'and'	<i>thuit e ceud troigh 's e a' streap air Aonach Dubh</i> <i>he fell 100 feet and him climbing on Aonach Dubh</i>	3.3.3
11. Main/matrix clause	<i>tha sinn a' falbh ...</i> <i>we are leaving ...</i>	3.2; 3.3
12. Relative clause	<i>an duine a thig gu tric</i> <i>the man who often comes</i>	3.3.2.3
13. Subordinate clause: cause or reason	<i>a chionn 's gun robh thu toilichte</i> <i>because you were pleased</i>	3.3.2.2
14. Subordinate clause: complement	<i>tha mi toilichte gun tàinig thu</i> <i>I am pleased that you came</i>	3.3.2.2
15. Subordinate clause: concessive	<i>ged a thàinig e</i> <i>although he came</i>	3.3.2.2
16. Subordinate clause: conditional	<i>ma thig thu</i> <i>if you come</i>	3.3.2.2
17. Subordinate clause: temporal	<i>nuair a thàinig thu</i> when you came	3.3.2.2

C. Verbal Morphosyntax

18. Aspect: perfect	tha sinn air sin a dhèanamh <i>we have done that</i>	2.3.1.2
19. Aspect: progressive	bha mi a' snàmh <i>I was swimming</i>	2.3.1.2
20. Aspect: prospective aspect	tha e gu tuiteam leis an acras <i>he is nearly collapsing with hunger</i>	2.3.1.2
21. Clausal negation	chan eil mi gad thuigsinn <i>I don't understand you</i>	2.3.1.3
22. Copular verb	's e, gur e, nach e, b' e, as, etc.	3.1
23. Copular construction: identificational clefts	's e fear-lagha a th' ann <i>he is a lawyer</i>	3.1.1
24. Mode: direct interrogative clause	a bheil thu a' tighinn? <i>are you coming?</i>	2.3.1.2
25. Mode: imperative verb	thalla! <i>leave!</i>	2.2.1.2
26. Tense: indefinite 1 ('Present- future')	bidh sinn ann a h-uile h-oidhche <i>we will be there every night</i>	2.3.1.2
27. Tense: indefinite 2 ('Habitual- past')	bhiomaid ann a h-uile h-oidhche <i>we would be there every night</i>	2.3.1.2
28. Tense: past definite	bha sinn ann reimhe <i>we were there before</i>	2.3.1.2
29. Tense: present definite tense	tha sinn ann a-nis <i>we are here now</i>	2.3.1.2
30. Valence decreasing: passive clause with agent	chaidh a dhèanamh aig Anna <i>it was done by Anna</i>	2.3.2.2
31. Valence decreasing: synthetic impersonals	thathar a' tuigsinn <i>it is understood</i>	2.3.2.2
32. Valence decreasing: total # of passive clauses	N/A	N/A

D. Nominal Morphology: Case, Adjectives and Appositives

33. Adjective: attributive	duine mòr <i>a big man</i>	2.2.4.1; 2.2.4.2
34. Adjective: demonstrative	an taigh seo <i>this house</i>	2.2.2.2
35. Adjective: predicative	tha an duine sean <i>the man is old</i>	2.2.4.2
36. Adjective: superlative/comparative	tha e nas fheàrr na mise <i>he is better than me</i>	2.2.4.4
37. Appositive	thuir Mo Mowlam rùnaire Èirinn a Tuath <i>Mo Mowlam, the Northern Ireland Secretary, said</i>	2.2.1.4
38. Complex noun phrase – 2 or more modifiers	seann taigh geal le doras uaine <i>an old white house with a yellow door</i>	N/A
39. Noun: genitive	ri taobh na mara(dh); ri taobh a' mhuir (sic.) <i>beside the sea (G.); beside the sea (N.)</i>	2.2.1.4
40. Noun: proportion of nonconcordant nouns to nouns overall	ri taobh a' mhuir (sic.), an fhear (sic.) <i>beside the sea; the man</i>	2.2.1.4
41. Nouns: total # of	N/A	N/A
42. Nouns: indefinite marking	taigh air choreigin; rudeigin <i>some house or another; something</i>	2.2.2.5

E. Prepositions, Pronouns and Deixis

43. Possession: with <i>aig</i> (periphrastic)	an càr agam <i>my car</i>	2.2.1.5
44. Possession: pronominal (simple)	a mhac <i>his son</i>	2.2.1.5
45. Prepositions: total	ri; airson; timcheall air <i>to; for</i> (complex prep.); <i>around</i> (compound prep.)	2.2.6
46. Pronoun: demonstrative	sin; seo <i>that; this</i>	2.2.2.2
47. Pronouns: total	N/A	2.2.2

F. Adverbs and Adverbials

48. Adverbs (other than those of place and time)	gu dòigheil; gu mall <i>fine; slowly</i>	2.2.5
49. Place adverbial	an-shin; an-seo; shuas <i>there; here; up</i>	2.2.5.3
50. Time adverbial	an-dràsta; a-màireach <i>now; tomorrow</i>	2.2.5.2
51. Adverbial: amplifier	<i>gives some indication of degree: e.g. gu mòr;</i> buileach; uamhasach <i>big; completely; terribly</i>	2.2.5.4
52. Adverbial: downtoner	<i>gives some indication of degree of uncertainty: e.g.</i> gu beagnaich <i>slightly</i>	2.2.5.4
53. Adverbial: emphatic	<i>only presence of certainty, no degree mentioned: gu</i> dearbh; gu deimhinne <i>indeed</i>	2.2.5.4
54. Adverbial: hedge	<i>less-specific indications of diminishing probability</i> <i>or uncertainty: e.g. cha mhòr; mu; 's dòcha etc.</i> <i>almost; about; perhaps</i>	2.2.5.4

G. Lexical Classes

55. Numeral	aon 'one'; an treas 'third'	2.2.3
56. Verb type: private	smaoineachadh; beachdachadh; creidsinn <i>thinking; considering; believing</i>	N/A
57. Verb type: public	ràdh; innse; mìneachadh <i>saying; telling; explaining</i>	N/A
58. Verb type: suasive	a' moladh; iarraidh, a' toirt air <i>suggesting; wanting; forcing to</i>	N/A

H. English Borrowing

59. Noun: assimilated	na compiutairean <i>the computers</i>	3.5
60. Noun: unassimilated	na computers	3.5
61. Words (non-nouns): unassimilated	duine <i>stupid</i> <i>a stupid man</i>	3.5

I. Lexical specificity

62. Average letters per word	N/A	N/A
63. Type-token ratio	N/A	N/A

4.3 Corpus Mark-up

Macaulay's (1991) work suggested that there were benefits to dividing the features into a clausal/discourse level group and one that was more concerned with the properties of individual lexical items. By making this distinction it was possible to treat clauses as the basic unit of syntactic analysis. They were defined as single verb structures having a

nucleus (the predicate) core, (the predicate and its arguments), and a periphery (see Appendix 1: §3.3 for more information). The first stage of mark-up involved dividing each text file into clauses via the insertion of hard returns; each 'paragraph' equaled one clause. After this, a two-column table was created and the text from the file was entered into the second column; each row equalled one clause. The clause-level tags were then inserted into column 1. After this was done, the word-level tags were entered into column 2 a-midst the text. The following is a small section of a tagged radio interview text:

Table 6: Example of a tagged text

FF CM	<a> còmhla rium <R2_PN> tha <VTI> Iain <NN> Caimbeul <NN>
CQ CM	Iain <NN> uell <LR> dè na <R> do <PPN> bheachdsa <MSDDC_LF>
FL CR CQ CM	an obair <FSDNC> a tha <VTI> ro <R> phàp' <MSIDC> ùr <AC> dè dh'fheumas <VFI> iad <PN> a bhith ' <SP> smaoin eachadh <VN_LV> air <RP>?
FF CM	 uell <LR> a's a' chiad dol-a-mach <LC> a Choinnich <NVC> tha <VTI> mi <PN> ga <SP_PPN> fhaighinn <VN> car <LD> mì- nàdarrach <AD> a' <SP> bruidhinn <VN_LP> air <R> cuideigin <FSIDC> a chuir <VL> an <R> àite <MSIDC> duine <MSIGC>

Key to Tags in Table 6

Clause level tags

CM	Main clause
CQ	Interrogative clause
CR	Relative clause
FF	Fronted adverbials and peripheral elements
FL	Left-detached elements

Word level tags

AC	Grammatically concordant attributive adjective
AD	Predicative adjective
FSDNC	Noun: feminine, singular, definite, nominative, grammatically concordant
FSIDC	Noun: feminine, singular, indefinite, dative, grammatically concordant
LC	Conjunct: <i>refers to clausal relations</i>
LD	Downtoner
LF	Focus clitic
LR	Discourse particle
MSDDC	Noun: masculine, singular, definite, dative, grammatically concordant
MSIDC	Noun: masculine, singular, indefinite, dative, grammatically concordant
MSIGC	Noun: masculine, singular, definite, genitive, grammatically concordant
NN	Proper noun: nominative (grammatically concordant)
NVC	Proper noun: vocative, grammatically concordant
PN	Pronoun
PPN	Possessive pronoun
R	Preposition
R2	Complex preposition
RP	Prepositional pronoun
SP	Progressive aspect
VFI	Verb: infinitivel ('future tense'), independent form
VL	Verb: infinitival complement
VN	Verb: verbal noun
VTI	Verb: definite present tense, independent form

At this point, the texts were converted into tab-delimited text files for compatibility to a custom-designed tag counting and concordancing programme: *LinguaStat*.

4.3.1 Issues and Limitations Posed by Mark-up

Although it was beneficial overall to be able to use two sets of tags in this way, there were cases where the tags were difficult or unwieldy in application. For instance, relative clauses in this scheme received the same treatment as subordinate clauses, although the former are a type of NP subordination (Van Valin and LaPolla 1997: 493). Scottish Gaelic has a high proliferation of relative clauses due to the commonality of cleft constructions and this resulted in parts of some texts looking rather jagged, and artificially few words per clause.

False starts, stutters, and other kinds of disfluencies posed problems for Macaulay (1991). In the current study, false starts that did not carry over to the following clause in any way (ex. *what did you ... did you go home after you left the party?*) were consigned to a line on their own and labelled with an appropriate clause level tag. Stutters called for a different treatment. Here, only one instance of each repeated word was tagged as seen in the following:

CO CN CM	agus <LO> cha robh iad cha robh <VPD> iad <PN> a' <SP> toirt <VN> dha <RP> biadh <MSIGU> no <LO> càil <PNI_MSIGC>
----------	---

Because the clause begins with *agus* 'and' and is coordinate, this is not a false start. In other words, part of the 'disfluency' is carried over to the 'fluent' part of the utterance, so it would be infelicitous to consign the first four words to a line labelled *false start*. Instead, the words were transcribed as spoken but only one repetition of them was actually tagged.

Where large parts of a clause were difficult or impossible to transcribe, or where the transcriber had been unsure of the transcription, no tags were used aside from one clause-level tag indicating 'incompleteness':

INC	a tha sinn a' dèanamh [?] gun smid aige
-----	---

Of course, there were times when it was not clear how individual lexical items should be tagged. This can be illustrated with English loans in ScG. A decision had to be made as to which should still be considered 'English' and which had been nativised into the language. With a word such as *buntàta*, having a long residence in ScG and having been assimilated to its phonology, it would be inappropriate to label it as a loan. However, in a case such as *prògram* 'programme', the proper path was not as clear. Decisions were made by consulting *Dwelly's Gaelic-English Dictionary* (1901–11) and Cox's Gaelic-Gaelic school dictionary (1991). If the item was not in either, it was considered to be a loan.

4.3.2 Grammatical Decisions: Adverbs and Case

A similar problem arose in determining the grammatical category of certain words, especially those of an 'adverbial' nature. The dictionaries are not always clear or consistent about such items and many had to be resolved on-the-spot given the features available in the tag-set.

Decisions concerning the case forms of a particular word were made, in the first instance, using the dictionaries mentioned. In Gaelic, there are a number of nouns with defective gender, appearing masculine in one condition and feminine in another, or subject to dialectal variation. Where a form was incongruent with the dictionaries, it was examined whether the speaker/writer was consistent in usage him/herself and whether the

form could be a dialectal variant. In some cases, items were determined to be non-standard but ultimately concordant. In other cases, they were tagged as nonconcordant.

4.4 Tag Counting

After all of the files were tagged completely, they were checked again manually for accuracy and to update the tags of the older texts. *LinguaStat* then counted all of the tags present (i.e. any text string in column 1 or any text string between '<' and '>' in column 2), normalising the frequencies to occurrences per 1000 words (see Figure 1 below).

Figure 1: KWIC-type Concordances in *LinguaStat*

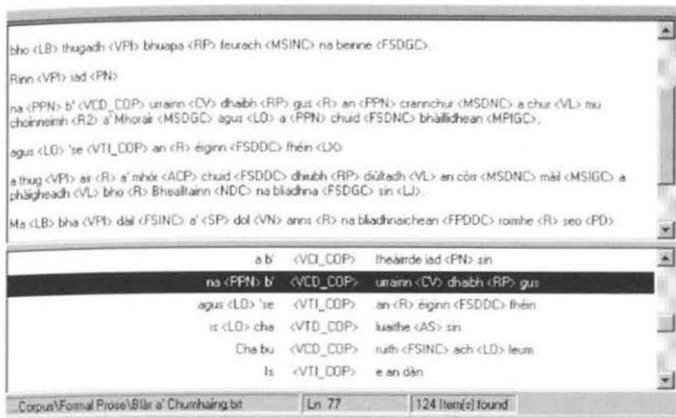
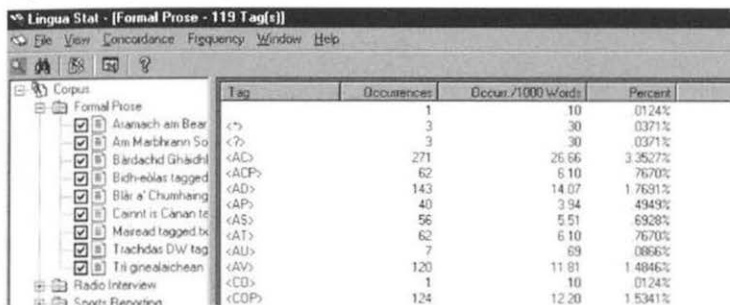


Figure 2: Tag Counting in *LinguaStat*



By exploring the frequencies and using the concordancing capabilities of *LinguaStat* (illustrated in Figures 1 and 2), a decision was made as to which features would be implicated in the quantitative analysis.

These features were then listed in a text file, in the form of individual searches with Boolean operators ('wild cards'). As mentioned previously, some of the features were combinations of individual tags. *LinguaStat* input this file to a macro which generated an Excel spreadsheet. One spreadsheet was generated per register. These were then combined into one master spreadsheet which was imported into SPSS 10.1 (Windows).

4.5 Data Analysis

Individual statistical tests were performed to establish whether there were significant differences between the eight registers in terms of each of the 63 features. An initial exploratory analysis determined whether the data met the assumptions of the chosen parametric procedure, the General Linear Model (GLM) Analysis of Variance (ANOVA): having (1) a normal distribution and (2) an equality of error variance in the dependent variable (feature frequency) across the different levels of the between-groups variable (the registers).

The Shapiro-Wilk test was used to measure distribution tendencies and was performed on each of the 63 individual data sets. Where a significant p value ($<.05$) obtained for more than half of eight registers in any particular set, a non-parametric test, Kruskal-Wallis, was used in place of the GLM ANOVA.³¹ This excluded a certain amount of data from the pairwise comparisons (see below), but it was still possible to consider the data and the trends that seemed to be represented therein.

Levene's test was employed to determine the level of heterogeneity in the within-groups variance. This test calculates the absolute difference between the value of each case in a group and the cell mean and then performs a one-way ANOVA on those differences. An exploratory analysis indicated that many of the individual comparisons would suffer from having unequal variance. To minimise the effect of this on the post-hoc analyses, Dunnett's T3 pairwise comparisons test was chosen, which is robust to departures from equal variance.

³¹ Overall, the GLM ANOVA is robust to departures from the normal distribution but it was considered important to have a cut-off point.

Introduction to the Results Chapters

The following three chapters present the quantitative and qualitative results of the study. As described in Chapter 4, sixty-three discrete linguistic features have been subjected to statistical analysis to determine their distribution in eight different registers. By understanding the fluctuations in the data in terms of the different contextual characteristics of the registers, it is possible to uncover the conventional associations in the corpus holding between context and linguistic form.

This is the first wide-scale, data-intensive study of register variation in Scottish Gaelic, or in any Celtic language. It incorporates several features which have been absent from or rarely used in past studies of register variation. In addition, this is the first time that many of the features appearing here have been investigated for Scottish Gaelic. This is especially true for those more characteristic of spoken rather than written language; almost all grammatical work on Scottish Gaelic up to the present time has concerned the latter.

The results are broken down into three chapters: Chapter 5 concerns features involving information structure and clause type, such as 'focus' constructions and subordination; Chapter 6 examines a wide range of features basic to Scottish Gaelic morphology and lexicon; Chapter 7 is an in-depth, quantitative and qualitative analysis of two features that were found to be robust register markers in the statistical analyses: nominal case concordance and noun phrase complexity. A profile of each of the individual registers based upon the findings of these three chapters can be found in Chapter 8. These profiles are important for reducing the data into generalisations, providing a means for examining the associations between groups of linguistic features and particular register types.

At the beginning of each results sub-section is a short introductory paragraph which details the characteristics of each feature, highlights its importance to the study, and surveys relevant past research. This is followed by a presentation of the data which provides means³² in tabular form and, where needed for interpretative purposes, additional descriptive statistics. Significant pairwise comparisons are either discussed in the text or illustrated with tables. A ball-ended arrow (●—●) is used to indicate statistically significant comparisons between two specific registers. Shading in the sorted data tables has been added to help visualise opposing groups or clusters of registers. The level of statistical significance for all tests was set at $p < .05$. It is important to note that, due to the effect of variance, some pairwise comparisons that may seem true based upon the means obtained are actually statistically non-significant. Each section contains commentary, wherein the significance of each finding is discussed and related to previous research and, where applicable, examples from the texts are provided. For more information on particular features, please refer to the chart and discussion in Chapter IV which lists, with examples, references to relevant pages and sections in the Scottish Gaelic grammar in Appendix 1.

The following is a list of the features most pertinent to ScG register variation, along with the contextual correlates to which they pertain. They have been sorted according to the order in which they appear in Chapters 5 and 6. The 'Degree' column states whether the feature is in positive ('>') or negative ('<') correlation with the contextual feature mentioned. To define some of the relevant contextual features that are used, **discourse freedom** is the degree to which a register evinces returning to previous points of reference or discursiveness. The feature **production constraints** concerns whether a register is produced on-line or with ample time for editing and consideration to linguistic form. A register is **informational** if it values information more than interaction or some other communicative function.

³² All means are per 1000 words of text.

<i>Feature</i>	<i>Section</i>	<i>Context/function</i>	<i>Correlation Direction</i>
demonstrative clefts	5.1	discourse freedom	>
left-detachment	5.2	production constraints	>
fronted adverbials/	5.3	discourse freedom	>
subordinate clause fronting	5.3	discourse freedom	>
clausal coordination	5.4.1	spoken monologue, narrative	>
clausal cosubordination	5.4.2	narrative	>
proportion of main clauses	5.4.3	informationality	<
temporal subordination	5.4.4.1	narrative	>
relative clauses	5.4.4.2	production constraints	<
present tense	6.1.2	oral non-narrative	>
past tense	6.1.2	narrative	>
indefinite 2 ('past-habitual')	6.1.2	narrative	>
total passives	6.1.3	informationality	>
clausal negation	6.1.4	informationality	<
interrogatives	6.1.5	interaction	>
imperatives	6.1.5	interaction	>
attributive adjectives	6.2.1	production constraints	<
demonstrative adj.'s	6.2.1	discourse freedom	>
complex NPs	6.2.3	informationality, production constraints	> <
total nouns	6.2.4.1	informationality	>
grammatical nonconcordance	6.2.4.2	production constraints	>
genitives	6.2.4.3	informationality	>
indefinite marking	6.2.4.4	interaction	>
prepositions	6.3.2	informationality	>
pronouns	6.3.3	narrative	>
private verbs	6.5.2	interaction	>
unassimilated English	6.5.3	production constraints	>
borrowing			
numerals	6.5.4	informationality	>
characters per word	6.6	informationality	>

The main findings that will emerge from Chapters 5 – 7 are as follows:

1. Despite being an endangered minority language, Scottish Gaelic exhibits many of the same patterns of register variation found in larger languages such as English
2. Conversational forms of the language diverge from other registers in specific ways regarding lexis, morphology, and syntax
3. There are differences between the registers in terms of English borrowing and grammatical conservativeness; some of these differences are related to formality level while others are due to production constraints and the attributes of the producers themselves
4. Registers that are more involved and/or interactive diverge significantly from those which are more detached and/or informationally oriented
5. Reportage-based registers differ from ones that are more discursive or imaginative
6. The use of devices for situating time varies according to whether registers frame events as sequential and dynamic (e.g. traditional narrative) or static (e.g. news scripts)
7. There are robust distinctions between narrative and non-narrative based registers

5 Information Structure and Clausal Types

This chapter investigates the way that prominence is given to certain constituents in Scottish Gaelic – referred to as information structure – as well as the distribution of a number of different clause types. In the section on information structure are examinations of the following: three types of clefts (i.e. *'s e* clefts, *'s ann* clefts, demonstrative clefts); left and right-detachment; and the placement of adverbials and subordinate clauses in relation to the matrix clause. The clausal section deals with the distributions and functions of the three different types of nexus (i.e. coordination, subordination, and cosubordination) and measures the proportion of main/matrix clauses in each of the registers, which provides a simple measure of clausal complexity. Finally, the chapter ends with a look at the use of asyndetic clausal juxtaposition in spontaneous spoken Gaelic, proposing that it assumes some of the same functions as subordinate clause combining in writing.

5.1 Clefts

Three main types of clefts were investigated as part of the quantitative analysis: *'s e* clefts; *'s ann* clefts (abbreviated forms of *is e* and *is ann* and closer to their usual pronunciation); and demonstrative clefts, being those headed by the definite deictics *seo*, *sin*, or *siud*. As discussed in the grammar (Appendix 1), *'s e* clefts are used to give emphasis to NPs and nominalised elements while *'s ann* clefts are used to highlight other kinds of constituents such as PPs, adjectives and adverbs. Both correspond roughly to IT clefts in English. Some examples follow, with the highlighted constituent in underlined type:

'S e cleft

- (1) *'s e prìs a' fuel a chuireas às do dhaoine an-seo co-dhiubh* (CONV)
'it's the price of fuel that will get rid of people here anyway'

'S ann cleft

- (2) *Gu dearbha, b' ann air èiginn – a bhriogais dheth 's am baidhseagail air a dhruim – a fhuair e nall gu Geàrraidh a' Bhota, dìreach mun do dhùin an fhadhail.* (FICT)
'Indeed, it was with difficulty – his trousers off of him and his bicycle on his back – that he made it to Geàrraidh a' Bhota, just before the ford closed.'

Reversed WH clefts (abbr. RWH), or demonstrative clefts *mutatis mutandis*, are described by Miller and Weinert (1998) as creating a bridge to preceding discourse or being used to pick out referents in the immediate temporal-spatial context. They say "RWH clefts regularly gather up previous discourse, clarify the present state of affairs and provide a starting point for a following piece of discourse, which often starts with an elaboration of the referent of THAT..." (p. 278). Overall, RWH clefts are characterised as performing less of a contrastive role than IT clefts. Example (3) follows of a demonstrative cleft from the Gaelic corpus:

Demonstrative cleft

- (3) *chuireadh [sic] ars' esan mo mhuime a-nuas mi a dh'iarraidh ars' esan na Cìre Buidhe ars' esan fhalbhaibh dha ars' ise siud i air an dreasair thall is thoir leat i* (TRAD)
'my mother he said sent me down to get he said the Yellow Comb he said go over to it she said that's it on the dresser over there and take it with you'

Miller and Weinert (1998) find that IT clefts are somewhat rare in spontaneous speech, but that RWH clefts tend to be more common. Biber *et al.* (1999) disagree, finding RWH clefts to be infrequent in all the registers they surveyed. Collins (1990) found that clefts in

general were more common in monologue and slightly more common overall in writing than in speech. Nelson (1997) observed no significant difference between clefts in speech and writing but found, like Collins, that they were more common in monologue, especially in scripted, public texts. In these registers, he states that they are used to organise discourse, as seen in introducing topics, changing topics, or setting out the context for the discussion of a topic. He also found that they occur with the highest frequency in spontaneous commentaries, such as on horse racing or football, where the foci of new and often contrastive information are constantly in flux. Finally, Biber *et al.* (1999) found that IT clefts were fairly common throughout English registers, but most abundant in academic prose. Demonstrative clefts, on the other hand, rarely occurred in academic prose but were found to be frequent in conversation.

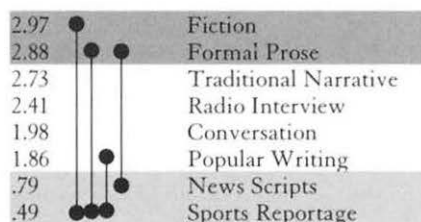
Data. Table 7 lists the means for these three features in the eight registers:

Table 7: Results: clefts

Register	'S <i>c</i>	'S <i>ann</i>	Dem <i>str.</i>
Conversation	2.14	1.98	2.57
Radio Interview	1.73	2.41	2.10
Sports Reportage	2.92	.49	1.32
Traditional Narrative	.63	2.73	2.10
SPEECH	1.86	1.90	2.02
Fiction	1.25	2.97	1.15
Formal Prose	1.40	2.88	1.20
News Scripts	.50	.79	.05
Popular Writing	1.98	1.86	.71
WRITING	1.28	2.03	.75
Total	1.57	1.97	1.38

A significant overall difference obtained for 's *ann* clefts ($F = 3.581$, $p = .002$). On a statistical level, this indicates that the group of registers selected exhibited feature frequencies that were more different from one another than one would have expected had they all derived from a single language variety. However, from this result alone, we do not know which individual registers are statistically different from one another. This information is provided by the pairwise comparisons, which are illustrated in Table 8 below:

Table 8: Sorted data: 's *ann* clefts



Examining the means on the left-side of Table 8, one sees that fiction and formal prose have the highest occurrence levels while sports reportage and news scripts have the lowest.

Each arrow indicates that a significant difference has been detected between the registers at its ends. Looking at the first arrow, we see a significant difference between fiction, is at the top of the scale for occurrence of *'s ann* clefts, and sports reportage, which is at the bottom. The second arrow shows that sports reportage and formal prose differed significantly. Overall, the distribution of arrows indicates a robust difference between fiction and formal prose on the one hand and news scripts and sports reportage on the other. The pairwise tests also detected a difference between sports reportage and popular writing. These findings will be interpreted below.

Demonstrative clefts were also found to vary significantly between the registers ($F = 4.95, p < .001$). The significant pairwise comparisons follow:

Table 9: Sorted data: demonstrative clefts

2.57	●	Conversation
2.10		Radio Interview
2.10	●	Traditional Narrative
1.32	●	Sports Reportage
1.20		Formal Prose
1.15		Fiction
.71		Popular Writing
.05	●●●	News Scripts

The distribution of arrows indicates a strong difference between the modes on this feature; all of the spoken registers are in the top half of the table while written registers are in the bottom half. Conversation had the highest return for demonstrative clefts while news scripts had the lowest. Significant contrasts were found between news scripts and three other registers (all spoken): conversation; traditional narrative; and sports reportage.

While a significant overall difference ($F = 3.006, p = .008$) obtained between the registers for *'s e* clefts, there were no significant pairwise comparisons. Because of this it is impossible to state which individual contrasts are statistically robust and which are due to chance. Perhaps the safest conclusion one could make is that sports reportage has more clefts than average in the corpus while news scripts contain less. Table 10 sorts the conversation data in descending order based upon occurrence level:

Table 10: Sorted data: *'s e* clefts

2.92	Sports Reportage
2.14	Conversation
1.98	Popular Writing
1.73	Radio Interview
1.40	Formal Prose
1.25	Fiction
.63	Traditional Narrative
.50	News Scripts

Commentary. Fiction and formal prose had the highest levels of *'s ann* clefts. The distribution of significant pairwise comparisons seems to indicate a distinction between registers with less editing and frequent changes of topic (news scripts and sports reportage), and those with more editing and less topic changes (fiction, formal prose, and popular writing). Perhaps even more importantly, the registers having more *'s ann* clefts are generally more discursive, while those having less tend to be, on the whole, more reportage-oriented. However, it is difficult to offer reasons for this particular distribution without additional work on the functions of *'s ann* clefts.

One observation may be offered regarding the difference between *'s ann* and *'s e* clefts: there is a semantic distinction between the deictics that are used in the two cases: *e*,³³ the third person masculine pronoun, simply identifies and points forward to the immediately following cleft constituent. This can be seen in example (4) from a conversation text:

- (4) *'s e Seonaidh Aileag [a chunnaic i]*
'it's Johnny Alick that saw her'

In Miller and Weinert's words (1998: 273), this type of cleft has identifying rather than predicating properties. *Ann* (> *is ann*/ *'s ann*), on the other hand – especially when followed by a prepositional, adjectival or adverbial constituent – often serves more of a predicating function; it indicates that a property is to be assigned to the proposition or entity in the cleft clause rather than the converse. An example of this from another conversation text is *'s ann car cruinn a bha i* 'it's somewhat round that she was'.

In opposition to previous research, neither the *'s e* nor *'s ann* cleft data show clear divisions on modal lines. However, the *'s e* cleft data does agree well with the findings of Nelson (1997: 344–5): in the corpus that he used – the one million word International Corpus of English – spontaneous commentaries had the highest level of cleft occurrences while news scripts had one of the lowest. Conversation also had a high return for *'s e* clefts. Admittedly, however, additional data is needed to bolster the statistical power of these comparisons.

Although modal differences were not detected in the other features, the demonstrative cleft data showed a robust distinction between speech and writing. The results differ somewhat from those reported in Biber *et al.* (1999): the news script register, rather than academic prose, was found to have the lowest occurrence of demonstrative clefts. The main contrasts were between conversation, traditional narrative, and sports reportage on the one hand, and news scripts on the other. Considering the strong forward motion of news scripts, with constant changes of topic, but little returning to previous mentioned referents, it is not surprising that there would be a paucity of demonstrative clefts. There would be little need for a device that reached up to previous discourse and managed the activation of contextually available referents. Conversation, on the other hand, is inextricably bound up with such requirements.

With high levels of both *'s e* clefts and demonstrative clefts, perhaps conversation, radio interview and sports reportage can be said to have more scope than most other registers. In this sense, scope is meant as the extent to which there are points of reference in the discourse or temporal-spatial context available for comment. A register with less scope, such as news scripts, is more confined in its movement between points of reference. Consequently, there is less of a need for constructions which serve to contrast points of reference or expand upon previously activated ones. Example (5) illustrates this presence of scope in conversation:

(5) **Conversation**

[2] *cha do reic e iad [na beathaichean] gus an d'ràinig e Carlisle*

[1] *tha fhios gum biodh iad air am fàs caol*

[2] *uell bha e a' falbh mun bhliadhna 'l fhios agad sin mar a bha iad a' deanamh bha iad ag obair fhios agad cha robh iad idir a' gabhail straight dhan an àite*

[2] he didn't sell them [the livestock] until he arrived at Carlisle

[1] no doubt they would have grown skinny [by then]

[2] well he travelled throughout the year you know **that's how they did** [it]
they worked you know they never went straight to the place

³³ Conservative usage would have *i*, the 3rd person singular feminine pronoun, when the cleft constituent has either feminine gender or reference.

Speaker 2 is discussing the transport of livestock in the old days and, in line 3 above, says a certain man used to travel throughout the year trying to sell his beasts. He then uses a demonstrative cleft to return to this statement – *sin mar a bha iad a' deanamh* 'that's how they did [it]' – followed by another statement which serves to sum up what he is trying to say. In contrast, a register such as news scripts rarely evinces this kind of returning to previous positions or contrasting different points of view. It progresses steadily from one topic to another, concentrating on relating facts, events, and the speech of others rather than dealing discursively with the subject matter. Example (6) illustrates this steady progression of topic, bounded by little except a brief pause.

(6) News script

Thuir Mo Mowlam gun deach an IRA glè fhaig air an fhois-fhòirneirt a bhriseadh leis a' phàirt a bh' aca ann am murt draibheir tacsaidh am Beul Feirste, ach nach do bhris i. Tha na Libearalaich Dheamocratach a' dol as àicheadh gu bheil iad ann am pòcaid a' Phàrtaidh Labaraich ann an co-bhanntachd Riaghaltas na h-Alba.

Mo Mowlam said that the IRA went very close to breaking the cease fire with their part in the murder of a taxi driver in Belfast, but they didn't break it. The Liberal Democrats are denying that they are in the pocket of the Labour Party in the coalition of the Scottish Executive.

To sum up the findings on clefts, in contrast to previous research no modal differences were found in the levels of 's *ann* and 's *e* clefts. However, 's *ann* clefts seemed to be more prominent in registers evincing higher levels of editing and more continuity of topic or discursiveness. It was suggested that the difference in the distribution of these two kinds of clefts, both of which would be translated as IT clefts in English, could be due to a difference in function, related in part to semantic distinctions between the deictics *e* and *ann*. The results for the demonstrative clefts showed a clear modal difference, with spoken registers having many more than written ones. In addition, registers with greater scope seemed to place a greater reliance on this construction while those with little scope had fewer instances of it.

5.2 Detachment

Two types of detachment were investigated: left and right. Detachment is a classic manifestation of spoken language's tendency towards syntactic fragmentation. One of its primary functions is to separate a referent from its role in a proposition, enabling easier processing by listeners. Lambrecht (1994: 185) says that there appears to be a general rule for spoken language: "do not introduce a referent and talk about it in the same clause". Prince (1997) avers that certain clausal positions are preferred over others for the introduction of new referents. Left-detachment simplifies the situation by introducing the referent outside the clause, so that it can be referred to within it using a pronoun. Biber *et al.* (1999) comment that left-detachment is used most commonly to establish topics, while right-detachment is frequently used to clarify or emphasise propositions. Miller and Weinert (1998: 238) maintain that left-detachment (or as they term it, an 'NP-clause' construction) is used primarily to highlight referents entering discourse as well as to help stage information in small chunks to ease processing. Previous researchers have found that left-detachment occurs much more readily in spontaneous spoken language (Miller and Weinert 1998; Biber *et al.* 1999) than other kinds of registers and that it is more likely to be found in lower-class than middle-class Scots speech (R. Macaulay 1991). Overall, this is an important feature for discerning spontaneous, informal speech from more formal, edited forms of discourse.

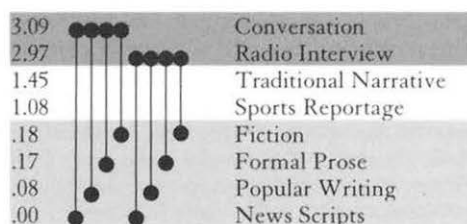
Data. Table 11 presents the findings for the Scottish Gaelic registers:

Table 11: Results: detachment

Register	Left	Right
Conversation	3.09	.99
Radio Interview	2.97	.00
Sports Reportage	1.08	.00
Traditional Narrative	1.45	.28
SPEECH	2.10	.34
Fiction	.18	.35
Formal Prose	.17	.00
News Scripts	.00	.00
Popular Writing	.08	.10
WRITING	.10	.10
Total	1.10	.22

There was a significant difference between the registers for left-detachment ($F = 11.566, p < .001$). The sorted data and pairwise comparisons are illustrated below:

Table 12: Sorted data: left-detachment



As the arrows show in Table 12 (each representing a single pairwise comparison), conversation and radio interview were significantly different from *all* of the written registers. These results are in keeping with previous research. The level of left-detachment in conversation and radio interview was very close. It was less prevalent in traditional narrative and sports reportage and returned very low levels in the written registers.

Right-detachment occurrences were tested using the Kruskal Wallis test and a significant difference obtained between the registers ($X^2 = 20.637; p = .004$). Conversation presented the highest levels of this feature, with instances also found in traditional narrative and fiction. Only two instances occurred in popular writing while there were none recorded in radio interview, sports reportage, formal prose or news scripts. Overall right-detachment was less common than left-detachment (also reported by R. Macaulay 1991).

Commentary. The results for both kinds of detachment indicate a much higher prevalence in spontaneous, conversational registers than any of the others. It is interesting to note that left-detachment was found at a relatively high level in all of the spoken registers but that right-detachment was, in fact, not found at all in 2 of the 4 spoken registers. Another interesting result was the presence of right-detachment in both narrative registers. Perhaps further research would show that it has a specialised function in these registers.

Miller and Weinert (1998: 238) comment that left-detached (or, as they call them 'NP-clause') constructions in spoken English are not necessarily the result of planning problems, as has sometimes been suggested. In the conversational Gaelic texts especially, the construction occurs with such regularity and with such a lack of the clutter associated with false starts that their comment seems applicable to this data as well. In addition, most of the noun phrases implicated in the construction are simple, indicating that its occurrence is not always motivated by the difficulties associated with heavy NPs. Consider the following examples:

- (7) *uell na Dòmhnallaich sin bha iad càirdeach rithe* (CONV)
'well those MacDonalds *they* were related to her'
- (8) *an duine aig a' chùl chì e na tha a' tachairt reimhe* (INTER)
'the man at the back *he* can see all that happens before him'
- (9) *linneagadh an taighe loisg iadsan loisg iad a h-uile sgath de linneagadh an taighe mu dheireadh* (CONV)
'the lining of the house *they* [EMPH] burnt *they* burnt every bit of the lining of the house finally'

In examples (7) and (8), a relatively simple extraclausal noun phrase occurs (*na Dòmhnallaich sin* and *an duine aig a' chùl* respectively) followed by a clause employing a pronoun which is coreferent with this noun phrase. Here we see the division of role and reference: the extraclausal noun phrase serves to establish the reference and the pronoun, which is informationally lighter, is used in the predicating or role establishing part of the utterance. Example (9) shows that, in Gaelic, the extraclausal NP can also be an emphatic pronominal. *Iadsan* is the emphatic form of the 3rd person masculine pronoun *iad*. This example is also interesting in that a listener might expect the first phrase *linneagadh an taighe* to be the extraclausal NP implicated in the detachment. However, as can be seen, this phrase, instead of being represented by a coreferent pronoun in the following clause, is repeated again. Example 3 shows the gradual build-up of information characteristic of spontaneous spoken language. In all of these examples, we simply have a 'topicalised' NP followed by a clause in which it is represented by a coreferent pronoun.

However, there are cases in the corpus where left-detaching does seem to be the product of planning difficulties, such as in example (10) from a conversation text:

- (10) Conversation example
- | | |
|---|-------------------------------|
| <i>'s bha e a' ràdh</i> | and he was saying |
| <i>gun robh màthair Eachainn;</i> | that Hector's mother, was |
| <i>gum faodadh i;</i> | that she, could give |
| <i>na thogradh i, thoirt dhi;</i> | what she, liked to her, |
| <i>nach do dh'fhàg i, dad riamh gun</i> | that she, never left anything |
| <i>phàigheadh</i> | without paying' |

The noun phrase *màthair Eachainn* is introduced as the subject of a complement clause but, immediately after uttering it, the speaker begins a new complement clause with a different verb (*faodadh* vs. *robh*). The first pronominal NP is not coreferential with *màthair Eachainn* but instead refers to a different previously mentioned referent. The coreferential pronoun occurs as an indirect object, in the form of a prepositional pronoun, *dhi*. In essence, not only did a left-detached construction occur, but it did so as part of a false start, an utterance that was not incorporated in following discourse.

Left-detachment regularly occurs with heavy noun phrases (see comments related to extraclausal NPs in §7.2). Although part of its motivation may lie in the processing difficulties associated with heavy noun phrases, it clearly has an identifying role as well. In example (11), again from a conversation text, the addresser offers a number of possible ways of identifying the referent in question, to be certain that the addressee knows whom she is talking about.

- (11) Conversation example 2
- | | |
|---|--------------------------------------|
| <i>cha chreid mi</i> | 'I don't believe |
| <i>nach cuala mi Dòmhnall Aileig</i> | that I didn't hear Donald Alick |
| <i>a' ràdh</i> | saying |
| <i>gun robh e am fear sin</i> | that he was that one of them |
| <i>dhiubh an t-athair do</i> | the father of the wife of |
| <i>bhean Thormaid an</i> | Norman the Blacksmith those |
| <i>Gobha na MacNeills sin</i> | MacNeills |
| <i>gur e an aon fhear</i> | that he was the only one |
| <i>a bha ceart a's an teaghlach aca</i> | who was trustworthy in their family' |

In line four, a complement clause is initiated with only a bare 3rd person masculine pronoun *e* as its subject. However, the addresser then includes, *am fear sin dhiubh* 'that one of them' and then, growing more specific, she offers *an t-athair do bhean Thormaid an Gobha* 'the father of the wife of Norman the Blacksmith'. At this point, it seems that the addressee has understood the reference, but just to be sure, she tags on *na MacNeills sin* 'those MacNeills'. After this phrase, matters of syntactic role are picked up again, and a new complement clause begins, but with a different verb – a dependent copular form *gur* rather than the dependent form of the past tense verb 'to be' *robh*.

As was mentioned, right-detached elements occur far less often in the corpus than left-detached ones. Their main function seems to be in clarifying or expanding upon the referent carried by a pronoun in the clause. In addition, they also share the same information spreading characteristic observed in left-detachment. A few examples will suffice to illustrate these points:

- (12) *och tha e cianail cearbach gnothach nam busaichean* (CONV)
'oh it is awfully clumsy the bus situation'
- (13) *"Agus tha coiseachd latha is bliadhna agad man ruig thu e," ars' esan, "do thaigh fhèin."* (TRAD)
"And you have a day and a year of walking until you reach it," he said, "your own house."
- (14) *B' iomadach clab a thug e leis, òrd a bha ùr o chionn dà bhliadhna.* (FICT)
'It was many a whack he gave it with it, a hammer that was new two years ago.'

In the first two examples, (12) and (13), the referent conveyed by the pronoun has not been explicitly mentioned in previous discourse. In example (12), although previous dialogue concerns a person's travails trying to cope with the Uist bus system, the third person pronoun could refer to a number of activated referents. The right-detached noun phrase establishes that it is the 'bus situation' which is 'awfully clumsy' rather than any of the other possible referents. In example (13), the intended referent of the pronoun *e* has not been mentioned *anywhere* in the previous discourse. In order to anchor this ambiguous anaphor, the extraclausal NP *do thaigh fhèin* 'your own house' is used. Finally, in example (14), although the referent is active and accessible – a 'hammer' was the topic of the previous two sentences – a right detached NP is used to expand its description. By separating this NP from a syntactic role, the writer can achieve a more concentrated focus of the aspect he wishes to highlight. One could say that it is language's equivalent of a stage spot light.

To summarise, left and right-detachment are found most frequently in spontaneous conversation. While left-detachment also occurs with regularity in radio interview, it showed no instances of right-detachment. Further research is needed to uncover the reasons for this discrepancy. Written registers had low occurrence levels of both forms of detachment but, notably, both narrative registers had instances of right-detachment where it was absent in some of the spoken registers. A number of examples were provided illustrating the main functions and the form of left and right-detachment.

Although previous research has characterised detachment as being a product of planning difficulties or, in Gaelic, as the breakdown of the verbal system (see Appendix 1: §3.4.2.2), it is suggested that one of the primary motivations for its use is the spreading of information out within discourse. To support this, a number of examples were advanced showing no false-start elements and the use of relatively simple extracausal NPs.

5.3 Placement of adverbials

This section covers two features: the fronting of adverbials and peripheral elements; the placement of subordinate clauses. These features have not been investigated in larger register studies (but see Biber *et al.* 1999). The fronting of an adverbial or peripheral element enables it to act as the theme, providing an entry into the following discourse via the information contained within it. It can be used for joining larger chunks of discourse or for navigating transitions. Biber *et al.* (1999) call these units 'linking adverbials' and says that they "can express a wide variety of relationships, including addition and enumeration, summation, apposition, result/inference, contrast/concession and transition" (p. 765). Chafe (1984a) says that expressions such as *moreover*, *however*, and *likewise*, when placed at the front of a clause matrix, have very similar functions to preposed subordinate clauses, serving "as a 'guidepost' to information flow" (p. 444). Some types of post-posed subordinate clauses, on the other hand, tend to behave like coordinate structures and have more of an afterthought quality to them than preposed ones. These attributes will be illustrated with examples in the interpretation sub-section below.

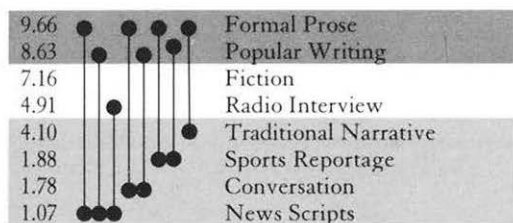
Data. Table 13 shows the occurrence data for fronted adverbials / peripheral modifiers:

Table 13: Results: fronted adverbials and peripheral modifiers

Register	Fronted adv.
Conversation	1.78
Radio Interview	4.91
Sports Reportage	1.88
Traditional Narrative	4.10
SPEECH	3.07
Fiction	7.16
Formal Prose	9.66
News Scripts	1.07
Popular Writing	8.63
WRITING	6.58
Total	4.83

A significant difference was found between the registers in occurrence of **adverbial and peripheral fronting** ($F = 10.483$, $p < .001$). Significant pairwise comparisons are illustrated in Table 14:

Table 14: Sorted data: fronted adverbials and peripheral modifiers



The table indicates that, for all of the written registers except news scripts, fronted adverbials and peripherals are fairly common. Formal prose and popular writing had the highest mean for the feature while news scripts and conversation had the lowest. The main distinction represented by the arrows is between formal prose and popular writing, on the one hand, and all of the spoken registers – aside from radio interview – and news scripts, on the other. This grouping is pointed out by the shading above.

In order to examine the **placement of subordinate clauses**, it is necessary to deal with ratios rather than strict occurrence levels. Table 15 displays the ratios of postposed and preposed subordination (in relation to the matrix clause) to instances of subordination overall:

Table 15: Results: placement of subordinate clauses

Register	Preposed Subord	Postposed Subord ¹⁴
Conversation	.33	.62
Radio Interview	.42	.57
Sports Reportage	.28	.72
Traditional Narrative	.48	.51
SPEECH	.38	.61
Fiction	.26	.72
Formal Prose	.27	.73
News Scripts	.07	.93
Popular Writing	.34	.66
WRITING	.24	.75
Total	.31	.68

Examining the proportion of preposed subordinate clauses, there was a significant difference between the groups ($F = 4.758$, $p < .001$). The pairwise comparisons showed that news scripts deviated from all other groups apart from sports reportage, conversation, and fiction yet other differences were not statistically significant. There was also a significant difference between speech and writing ($F = 9.851$, $p = .002$), yet it stands to reason that much of the variance in this comparison was taken up by news scripts alone.

Overall, these findings indicate that there is a preference, in most registers, towards postposed subordinate clauses rather than preposed ones. Traditional narrative and radio interview seem to be the only registers with a more even distribution of preposed and postposed subordinate clauses.

Commentary. It was seen that **adverbial fronting** was common for all of the written registers except news scripts. This implies that it is more a feature of discursive rather than reportage-based language. Additionally, it was fairly prevalent in radio interview and traditional narrative, but not in the other two spoken registers. The common function of this feature is to manage transitions of various kinds between stretches of discourse, or to situate a following clause in space, time, or manner. The following examples demonstrate some of these characteristics:

- (15) *Aig amannan bidh mi faireachdainn aonaranach.* (FICT)
'At times I feel lonely.'

¹⁴ Due to a few cases of free subordinate clauses (those present independent of a matrix clause) in most of the registers, the row totals do not always add up to 1.

- (16) *Gun ghuth ri chèile, sheas Tormod is a bhan-chompanach fhèin taobh ri taobh.*
(FICT)
'Without a word to each other, Norman and his own female companion stood side by side'
- (17) *'s ann leam iad uileag arsa Fear a' Churracain Ghlais agus bidh do cheann agam còmhla riuth' agus leis a-sin shad e 'n ceann dheth* (TRAD)
'they are all mine said 'The Man of the Grey Cap' and your head with be mine along with them and with that, he struck his head off'
- (18) *'S ann air bàrdachd de'n dàrna seòrsa sin a tha an comh-chruinneachadh sa a-mach. An dèidh sin cha bu mhisde sinn tòiseachadh le sùil ghoirid a thoirt air cruth agus ionad na bàrdachd ghnàthasaich* (FORM)
'It is with poetry of the second type that this collection is concerned. Despite that it would do us no harm to begin with a quick look at the form and context of traditional poetry.'
- (19) *agus tha sinn a' coimhead air rud an-sheo a bhiodh maireannach ach air an làimh eile chan eil fhios agamsa an e sin an dòigh a b' fheàrr air a chur air chois* (INTER)
'and we are looking at a thing here that would last but on the other hand I don't know if that is the best way of establishing it'

Example (15) above does not manage a transition, but rather situates the following clause in a temporal context: *aig amannan* 'at times'. Example (16) thematises the manner in which the participants of the following clause complete the act described. Example (17), with *leis a-sin* 'with that', incorporates both a temporal transition and summarising function. The proposition after this fronted element (him striking off his head) is cast as occurring right after the speech event of the previous clauses. The demonstrative pronoun is used to refer back to this event. The reading of summation or finality is also present – he struck off his head and 'that was that'. Example (18) has a linking adverb associated with concession, *an dèidh sin* 'despite that' [Lit. 'after that'], while example (19) shows a transition of contrast: *air an làimh eile* 'on the other hand'. Both examples make explicit references to previous propositions.

Returning to the results, perhaps the reason that radio interview has such a high level of the feature for a spoken register lies in the nature of interview. An onus is upon the interviewer to sum-up and recast previous discourse, trying to gain additional angles on the topic of discussion. Both popular writing and formal prose tend to be discursive registers, involved with examining different viewpoints and, it appears, having a need to explicitly mark links between propositions. Although fiction and traditional narrative do not have this same need, they place a heavy requirement on situating time and place as the narrative progresses. A large proportion of examples in these two sub-corpora evince temporal transition management. Neither sports reportage, conversation, nor news scripts appear to need regular, explicit marking of the kinds of relationships in evidence between clauses in the other sub-corpora. In news scripts, this may be able to be explained by the general absence of references to previous discourse. In conversation, it may be that juxtaposition alone suffices (see §5.4.5). Finally, sports reportage does not generally deal in discussion and there is an iconic relationship between the temporal progression of the game and the language used to describe it; events are reported as they happen and there is scant need for explicit temporal linking of propositions.

The results for subordinate clause placement indicate that most registers had a clear preference for these clauses to occur after a matrix clause. News scripts, in particular, very rarely had subordinate clauses in a fronted position. Once again, this is probably due to the strong forward inertia of the register, with little revisiting of previous

discourse in evidence. Traditional narrative and radio interview were the only registers to exhibit tendencies in the other direction. The following examples are typical of these three registers:

- (20) *nuair a bhiodh e 'cuir a-staigh na cruaiseachd bhiodh e 'cuir a' bhèin a' falach fo na sparran gum biodh e 'deanamh na h-ath-chruach an ath-bhliadhna* (TRAD)
'when he was putting in the haystack he would hide the skin beneath the cross-beams until he did the next haystack the following year'
- (21) *can ma thig an ola is ma gheibh iad ola a lorg a-muigh air taobh a-tuath shuas cùl a Hiort an-shin tha am port-adhair ann am Beinn na Faoghla gu math nas fhaise air na tha am port-adhair ann an Steòrnabhagh* (INTER)
'say if the oil comes and if we manage to find oil out on the north side up behind St. Kilda there the airport in Benbecula is much closer to it than the airport in Stornoway'
- (22) *Bha dithis a bhuinneadh do theaghlachan Hiortach a-measg ma dhà fhichead a thàinig cruinn airson na tha 'n dàn dhan eilean a dheasbad nuair a chaidh pàrlamaid an eilein a ghairm airson an latha 'son a' chiad turas ann an cha mhòr trì fichead bliadhna 's a deich* (NEWS)
'There were two people from St. Kildan families amongst the about forty who gathered round to debate the fate of the island when the Parliament of the Island was called for the day for the first time in almost seventy years'

In traditional narrative (seen in example (20) above), by far the largest percentage of fronted subordination comes from adverbial clauses of time. Most of these are initiated with the subordinator *nuair* 'when'. This register places a heavy need on text-internal reference of time to situate occurrences and relate them to each other in the narrative. Radio interview on the other hand, while having relatively few temporal clauses overall, frequently employs fronted conditional clauses. The main function of a conditional clause, illustrated in example (21) above, is to state the truth conditions for the proposition contained in another clause. By occurring in a fronted position, it establishes relevance to previous discourse as well as providing an entry point into following discourse. This type of intermediate positioning between previous and following information is largely absent in news scripts. Most subordinate clauses in this register occur after the matrix clause, as seen in example (22) above. The difference between this use of temporal clauses, and that of traditional narrative, can be largely attributed to the discontinuous nature of news scripts, which has frequent topic changes and little back movement, therefore less need of what Chafe (1987) referred to as 'guideposts'. This is perhaps the most robust difference, both functionally and structurally, in evidence in the comparison between the registers on this feature.

5.4 Clause types

5.4.1 Coordinate Clauses

Spoken registers have been reported as having much higher levels of coordination than written ones (Kroll 1977; Chafe and Danielewicz 1987; Miller and Weinert 1998). Conversation in particular, and spontaneous speech in general, is characterised as having frequent instances of coordination in the form of simple joined strings (Chafe and Danielewicz 1987; Greenbaum and Nelson 1995; Miller and Weinert 1998). However, Greenbaum and Nelson state that, overall, conversation evinced lower levels of

coordination compared to unscripted monologue and radio broadcast. They interpret the higher amount in broadcasts, despite being interactive like conversation, as being due to the presence of an interviewer and possibly a difference in planning time and formality.

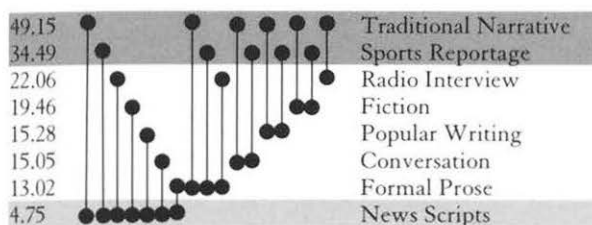
Data. Table 16 presents the means for the occurrence of coordinate clauses found in the Gaelic corpus:

Table 16: Results: coordinate clauses

Register	Coordination
Conversation	15.05
Radio Interview	22.06
Sports Reportage	34.49
Traditional Narrative	49.15
SPEECH	30.61
Fiction	19.46
Formal Prose	13.02
News Scripts	4.75
Popular Writing	15.28
WRITING	12.85
Total	21.73

A significant difference obtained between the registers for coordination ($F = 29.754$, $p < .001$). Pairwise comparisons showed a wide range of significant contrasts as illustrated below:

Table 17: Sorted data: coordinate clauses



In essence, there was a three-tiered division of the data. News scripts were at the bottom, distinct from all other registers. In the middle tier were formal prose, conversation, radio interview, fiction and popular writing. Finally sports reportage and traditional narrative had a significantly higher occurrence level of coordination than the other registers. The most divergent contrast was between traditional narrative and news scripts.

Commentary. These results differ somewhat from those of other researchers. Perhaps this is due in part to the inclusion in this study of certain registers that were absent in others, such as traditional narrative and sports reportage. If one looks closely at the results, it is possible to see additional reasons. In traditional narrative, which has the highest level by far of coordinate clauses in the corpus, communication seems to take the form of long bursts of description and dialogue, iconically occurring in temporal succession. These bursts are linked together, mainly, using the words for 'and' – *agus* or '*s*' – as seen in example (23), which is typical of this text type, from the narrative *Garaidh agus na Fiantaichean* 'Garaidh and the Fenians':

(23) Example from Traditional Narrative illustrating coordinate clauses

bha taigh aig na boirionnaich dhaibh péin, agus bha Garaidh [a'] watch-adh nam boirionneach agus nuair a chlioradh na Fiantaichean air falbh, sgioblaich iadsan iad fhèin agus chaidh iad dhan tràigh. Agus thàinig a h-uile gin aca agus sgùird dhe mhaorach aca³⁵ agus bhruidh iad e ...

the women had their own house, and Gary was watching the women and when the Fenians cleared off, they tidied themselves up and they went to the shore. And every one of them came and a lap-full of shellfish at them ['and they had a lap-full of shellfish'] and they cooked it ...

As is apparent, *agus* occurs after every clause save one. Ostensibly, this use of coordination is not only for the stringing together of linked clauses, but also for maintaining the pace of the story in question, and, therefore the attention of the audience. In addition, like certain other fragmentary devices such as left-detachment, it may aid listeners' comprehension by carving utterances into small, easily understood chunks. Finally, unlike dialogue, which is punctuated by interruptions and turn taking, traditional narrative runs its own course and may benefit from a more explicit demarcation of language flow to cope with human processing constraints.

Sports reportage is, in some ways very similar to traditional narrative: it deals in descriptions bound to a relatively fixed temporal flow; it requires some way of demarcating the course of the language; and it is a largely monologic register. Example (24) illustrates dense coordination in a sports reportage text:

(24) Example of dense coordination

Gough ga feuchainn suas an loidhne gu Nicol ach choinnich Spasitch i agus dh'fheuch esan gu Sretgovitch i agus tha an rèitire ag ràdh gun do rinn Spasitch droch chluich an-siud air Nicol thàinig Nicol a-nuas agus 'se buille shaor a th' ann do dh'Alba

Gough trying it up the line to Nichol but Spasitch met it and he tried it to Sretgovitch and the referee is saying that Spasitch made a bad play there on Nicol Nicol came up and it's a free ball for Scotland.

Example (24) looks very similar to the traditional narrative one, with a coordinator occurring after almost every clause. When dense coordination like this occurs, the functions seem very similar to those mentioned above. However, overall, description occurs differently in sports reportage. Most plays are reported in the form of asyndetic chunks, with often only a very short pause or a name shift to indicate that a new chunk is beginning. This is the juxtaposition that Miller and Weinert (1998) report as being basic to spontaneous spoken language. Explicit coordination seems to be reserved, in general, for more pivotal sequences of action, such as when a ball has been stolen, as above, or when players are advancing upon the goal. It seems to say "listen, this is important". Further points and examples related to this use of coordination can be found in section 8.3 'Sports Reportage'.

Conversation had the second lowest return for coordination, which contrasts to some previous research, but not Greenbaum and Nelson (1995). After looking closer at the current data, it was found that the one text with the highest occurrence of coordinated clauses (24.02 per 1000 words) was the only 2-person dialogue represented in the corpus; all other conversations had more than 2 participants present (average -4.5). There may be differences in the linguistic form of conversations as the number of participants increases.

³⁵ This particular clause, being non-finite, is cosubordinate rather than coordinate.

With less participants, turns can be longer, allowing more room for extended narrative, which seems in this corpus to correlate with higher levels of clausal coordination.

5.4.2 Cosubordinate Clauses

Cosubordinate clauses in Scottish Gaelic are formed when a matrix clause, using the conjunction *agus*, joins with a subsequent nonfinite clause. Example (25) illustrates this construction:

- (25) *'s bha 'm boireannach a' tighinn às a dhèidh 's i 'rànaich ag iarraidh a' bhèin*
(TRAD)
'and the woman was coming after him and her crying wanting the skin'

The important characteristic of the clause in bold is that it does not contain a finite verb. Because of this, it is brought semantically closer to its antecedent clause when conjunction occurs between them. Clausal cosubordination of this type is somewhat unique to Gaelic and as far as I am aware has not been studied previously within the context of register variation. For more information on the construction and further examples, see Appendix 1: §3.3.3.

Data. A significant difference was found in the relative prevalence of clausal subordination in the registers ($F = 4.266$, $p < .001$), yet the only statistically significant post-hoc comparison (Dunnett's T3) was between radio interview and news scripts. However, this does not seem to tell the whole story if other measures are considered.

Table 18: Results: cosubordinate clauses

Register	Mean	Median	Minimum	Maximum	Std. Deviation
Conversation	.84	.00	.00	4.27	1.39
Radio Interview	.67	.83	.00	1.83	.64
Sports Reportage	1.26	.96	.00	3.96	1.16
Traditional Narrative	2.60	2.72	.00	5.02	1.69
SPEECH	1.38	.95	.00	5.02	1.47
Fiction	4.64	2.42	.96	13.44	4.51
Formal Prose	1.77	.99	.00	4.66	1.61
News Scripts	2.86	2.79	.98	5.29	1.52
Popular Writing	.82	.00	.00	4.56	1.36
WRITING	2.38	1.47	.00	13.44	2.73
Total	1.88	.99	.00	13.44	2.24

Commentary. If reference is made to the medians – which seem to be more appropriately consulted in this data due to the effect of outliers and large amounts of variance present in some registers (especially fiction and formal prose) – it is apparent that the overriding discourse type displaying the highest levels of cosubordination is narrative, although a very high level obtained in news scripts as well. With more data, perhaps this division would become even more clearly defined.

One might be tempted to question how this clause type is any different from general coordination, albeit presenting ellipsis of the finite verb. First, although it incorporates the conjunction *agus* it clearly contrasts in distribution to coordinate clauses, as covered in §5.4.1: the news script register has one of the highest frequencies of *agus* cosubordination but one of the lowest of coordination; also fiction, while having a high occurrence level of cosubordination, had an average level of coordination. Second, as covered in §3.3.3 Appendix 1, the interpretation of expressions of this type is often very different from what would obtain with standard coordination. One example will suffice to illustrate this point:

- (26) *Eil fhios agad, uair no dhà, 's **mi fighe air a' bheairt sin thall**, shaoil mi gu robh e agam, cha mhòr. (FICT)*
 D'you know, once or twice, **while I [was] knitting on that loom over there**, I thought that I had it, just about.

If the cosubordinate clause in example (26) [in bold type] is provided with a main verb, perhaps *bha* 'was', the interpretation would alter significantly:

- (27) *Eil fhios agad, uair no dhà, 's **bha mi fighe air a' bheairt sin thall**, shaoil mi gu robh e agam, cha mhòr. (FICT)*
 D'you know, once or twice, **and I was knitting on that loom over there**, I thought that I had it, almost.

Instead of indicating cotemporality and a close semantic link with the following matrix clause, as in the cosubordinate example, the coordinate clause is interpreted as more of a parenthetical remark. Example (27) has the reading of 'by the way' or 'as a matter of fact'.

Agus cosubordination can carry a number of different interpretations apart from cotemporality (see Appendix 1) and more research is needed on its discourse functions in a range of different types of text. This would constitute a study in its own right, however, and available space and time preclude such an undertaking at the present time.

5.4.3 Main/matrix Clauses

The proportion of main or matrix clauses found in a text gives some indication of its level of clausal complexity. The higher the concentration found, the more probability there is that the text tends towards simpler clause structures or, more specifically, that the main clauses found in the text tend to be shorter, with fewer lexical items.

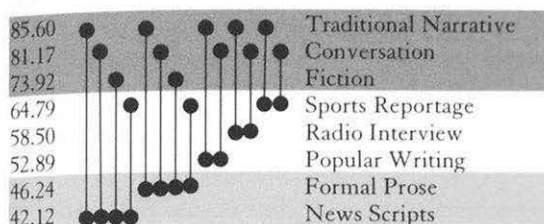
Data. There was a significant difference between the registers in main/matrix clauses ($F = 24.931, p < .001$).

Table 19: Results: main/matrix clauses

Register	Mean
Conversation	81.17
Radio Interview	58.50
Sports Reportage	64.79
Traditional Narrative	85.60
SPEECH	73.26
Fiction	73.92
Formal Prose	46.24
News Scripts	42.12
Popular Writing	52.89
WRITING	52.91
Total	63.08

There were also a large number of significant pairwise comparisons as seen in Table 20:

Table 20: Sorted data: main/matrix clauses



The arrows indicate a major contrast between news scripts and formal prose, on the lower end of the scale, and fiction, conversation and traditional narrative on the higher end. As is clear from the means on the left side of the table, there was a wide range of variation between the registers for number of main clauses, a difference of over 43 occurrences per 1000 words between news scripts and traditional narrative. There was also a large difference between the modes for the feature, seen in the means for speech and writing above.

Commentary. Clearly, traditional narrative and conversation have the highest proportion of main or matrix clauses in the register sample. This finding indicates that they tend to have smaller, simpler clauses than many of the other registers. At the low end of the list are formal prose and news scripts which, it seems, have a tendency for either more clause combinations and/or longer main or matrix clauses than the other registers. If clauses had not been defined by the presence of a finite verb, sports reportage would have scored much higher on this feature. As was illustrated in §5.4.1 above, many of the utterances in sports reportage are simply sequences of referents and prepositions, without a main verb present at all. Overall, a high incidence of clause combining indicates more planned language production and is one of the characteristics of 'integrated' language. Generally speaking, those registers with a higher number of main clauses also tend to have other 'fragmented' features such as more left-detachment, fewer attributive adjectives, and fewer complex noun phrases.

The following examples illustrate the difference between the two most opposing registers on this feature, traditional narrative and news scripts. The examples are divided into 10 clauses each. The important thing to note is that the news script excerpt (example (29)) is considerably more dense than the traditional narrative one (example (28)); while the news script has 113 words, the traditional narrative example only has 60:

(28) Traditional Narrative

1. *bha Mac Iain 'ic Sheumais air toll*
2. *a bh' air a' bhalla dìreach mu choinneamh toll*
3. *a bh' air a' bhalla*
4. *agus saighead aige*
5. *agus thuirt e*
6. *mheall mo làmh mi ars' esan*
7. *neo cha dèan thu sin*
8. *agus dh'fheuch e saighead air an duin'*
9. *agus chuir e dìreach ann an ubhal a' sgòrnain aige i*
10. *agus bhrùchd càch a-mach*

(29) News Script

1. *Tha gille beag ceithir bliadhna a dh'aois a' tighinn air adhart 'san ospadal am Bruach Chluaidh as dèidh ionnsaigh a thoirt air le cù.*
2. *Chaidh Seumas Longridge a ghortachadh mun aodann 's mun bhodhaig*
3. *Chaidh an cù – Japanese Akita – a chur gu bàs le lighiche-sprèidh.*
4. *Tha aithisg air a' chùis a' dol gu neach-casaid a' Chrùin.*
5. *Tadhlaidh oifigearan Comhairle nan Eilean air dà choimhearsnachd ann an Nis an Leòdhas*
6. *a tha 'g iarraidh cidhe agus caladh a leasachadh.*
7. *Cuiridh Comhairle Choimhearsnachd Nis coinneamh air dòigh ann an Sgoil Lionail air an t-siathamh-latha-deug dhen ath mhìos*
8. *gus an cluinn iad*
9. *na tha dhìth air muinntir an àite.*
10. *Ron a' choinneamh chì iad Port Nis ...*

5.4.4 Subordinate Clauses

Various types of subordinate clauses were surveyed for this research: relative clauses; cause/reason clauses; complement clauses; concessive clauses; conditionals; and temporal (adverbial) clauses. Some returned relatively high occurrence levels (temporal subordination and relative clauses) while others were so infrequent as to be inconclusive (cause and reason clauses and concessive subordination). In general, previous research has indicated that subordination is more common in writing than speech (see references in Chapter II). The current findings indicate that, while this seems to be true for the corpus, the individual types of subordination vary in their distribution. Miller and Weinert (1998) account for some of the tendency for written language to have more subordinate clauses by explaining that spontaneous spoken language often avoids it through stringing main clauses together *asyndetically*. This topic is explored using Scottish Gaelic data in §5.4.5 below, with examples from the conversation sub-corpus.

5.4.4.1 Temporal Subordination

A number of different subordinators are used in ScG for temporal adverbial clauses, the most common in the corpus being: *an uair/nuair*; *aon uair 's gun*; *am feadh*; *(bh)o*; *cho fad's*; *gun/gum/gus*; *mun/mum/mus*; and *on*. Biber (1995) mentions that temporal subordination occurs more frequently in narrative, but this feature has received scant attention in other register studies.

Data. A significant contrast was found between the registers ($F = 8.869$, $p < .001$).

Table 21: Results: temporal subordination

Register	Mean	Median
Conversation	3.15	2.73
Radio Interview	2.89	2.74
Sports Reportage	1.14	.97
Traditional Narrative	9.34	9.09
SPEECH	4.20	2.99
Fiction	7.77	6.36
Formal Prose	2.79	2.71
News Scripts	2.62	2.19
Popular Writing	4.24	3.93
WRITING	4.21	3.50
Total	4.21	3.03

Several pairwise contrasts were significant as well, as illustrated below:

Table 22: Sorted data: temporal subordination

9.34	●	Traditional Narrative
7.77	●	Fiction
4.24	●	Popular Writing
3.15	●	Conversation
2.89	●	Radio Interview
2.79	●	Formal Prose
2.62	●	News Scripts
1.14	●	Sports Reportage

The arrows indicate a strong contrast between traditional narrative, which had the highest occurrence of the feature, and all of the other registers except popular writing and fiction. There was also a significant difference between sports reportage, which had the lowest level of occurrence, and popular writing. Considering the range of the means, one can see that temporal subordination occurs more than eight times more frequently in traditional narrative than it does in sports reportage.

Commentary. The data concur well with Biber's finding: traditional narrative has the highest occurrence of temporal subordination, followed by fiction and then popular writing. Traditional narrative and fiction in particular rely heavily on text-internal temporal reference for the placing of propositions in relation to one another. Narrative registers are concerned with unfolding events. These are not features of news scripts, where events are presented as static, nor of sports reportage, where time reference is implicit and bound to the present; discussion of other time frames does not usually occur. Additionally, as was discussed in §5.4.3, sports reportage and news scripts, because of their fact-based nature and strong forward inertia, do not evince frequent clause linking.

Example (30), from traditional narrative, shows the way that events are presented as unfolding in this register:

(30) Traditional narrative example

<i>dh'fhalbh an gille a-null</i>	the boy went over
<i>'s a' chìr air an dreasair</i>	and the comb on the dresser
<i>'s e a' dol a-null dha togail</i>	and him going over to lift it
<i>'s mun do ràinig e an dreasair</i>	and before he reached the dresser
<i>thog ise slacan draoidheachd</i>	she [a witch] raised a magical wand
<i>agus bhuail i air a' ghille e</i>	and she hit the boy with it
<i>agus rinn i fitheach dubh dhe</i>	and she made a black raven out of him
<i>Agus nuair a bha e a' dol a-mach</i>	And when he was going out the door
<i>an doras</i>	
<i>thug e aghaidh air an doras</i>	he went to leave through the door
<i>agus nuair a bha e a' dol a-mach an</i>	and when he was going out the door
<i>doras</i>	
<i>chuir e bruchd faladh a-mach</i>	he spat out a drop of blood

Nearly every clause in this sample encodes a different event or propositional frame, progressing in temporal succession: the boy goes over; the comb is on the dresser; he is going to lift it; before he reaches it; the witch raises a magic wand, and so on. In addition, there are three instances of temporal subordination (one is repeated), in bold type, which serve to situate one proposition within the time frame of another. In news scripts, time references are usually presented as static, through the use of temporal adverbs. As detailed in §6.4, the news script register (followed by sports reportage) actually has the highest significant occurrence of temporal adverbs. We can see in example (31) how these two features differ functionally:

(31) News script example

*Tha dùil
gun cùm ceathrar Bhreatannach
a chaidh a thoirt am brùid ann an Liberia
coinneamh naidheachd aig port-adhair
Heathrow an ceart-uair
agus iad air ùr thilleadh madainn
an-diugh.
Thathas ag aithris
gu bheil iad gu math – ach am feum
beagan fois.*

It is expected
that four Britons -
who were taken hostage in Liberia
- will be keeping a press meeting
at Heathrow just now
with them having just returned this
morning.
It is reported
that they are well – but needing a
little rest.

Instead of relating events or propositional frames to each other, news scripts are concerned with expressing the exact time of occurrences. In example (31), time is coded by the adverbs *an ceart-uair* and *madainn an-diugh*. There are no cases of clause combinations to establish temporal reference. While temporal subordination provides text-internal context – stabilised through a chain of propositions in relation to one another – temporal adverbs make reference to static, text-independent points of reference (today, tomorrow, yesterday, right now, etc.) to situate occurrences.

5.4.4.2 Relative Clauses

Relative clauses are a type of NP modification capable of both identifying a noun and extending its semantic sense. In the current study, they could be either headless (incorporating the headless pronouns *far* or *na*) or headed, in which case the relative pronoun *a* was usually present. The only time that a relative pronoun would not have been present in the Scottish Gaelic data would have been in the case of pied-piped prepositional relatives such as *tha creag mhòr [ris an cainnte]* 'there is a big rock [to which would be called]'.

Previous studies have indicated that, in unplanned discourse, referents are more often marked using determiner plus noun constructions (e.g. *this man ...*) than full relative clauses (Ochs 1979). Additionally, left detachment can appear in unplanned discourse where relative clauses would be found in more planned language styles (Ochs 1979; also see section 5.4.5 on juxtaposition in spoken language). Chafe (1984) found that restrictive relative clauses were twice as common in writing *versus* speech. For every 1000 words of text, he located 15.8 relative clauses in writing as opposed to 9.7 in speech. Miller and Weinert (1998) find less variety in the types of relative clauses used in spontaneous spoken language as opposed to other language varieties. For other studies see: Beaman (1984); Biber (1988); Fox and Thomson (1990); and Kirk (1997).

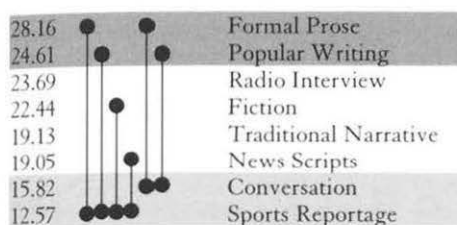
Data. There was a significant main effect for relative clauses ($F = 7.61, p < .001$) and register. The difference between speech and writing was also investigated for this feature and found to be significant ($F = 15.971, p < .001$).

Table 23: Results: relative clauses

Register	Mean
Conversation	15.82
Radio Interview	23.69
Sports Reportage	12.57
Traditional Narrative	19.13
SPEECH	17.49
Fiction	22.44
Formal Prose	28.16
News Scripts	19.05
Popular Writing	24.61
WRITING	23.53
Total	20.51

Pairwise comparisons are illustrated below:

Table 24: Sorted data: relative clauses



Relative clauses are most prevalent in formal prose and popular writing while they are relatively infrequent in conversation and sports reportage. The main contrasts, as indicated by the arrows and shading, are between these two groups. Other significant contrasts were sports reportage *versus* fiction and news scripts.

Commentary. It is not clear in other studies whether or not the relative clauses that were tagged included those implicated in clefts. In Scottish Gaelic, cleft-type constructions are common due to the heavy loading on them for identificational and highlighting constructions (see Appendix 1). Assuming that Chafe (1984) did not include clefts in his relative clause data, the current results are similar to his own. As a general trend, more edited and literary registers have a greater occurrence of relative clauses. Radio interview is an exception to this but, as it did not differ significantly with any other register, we cannot be sure of its placement in relation to them. As mentioned previously, the functions of relative clauses are sometimes managed in spoken language through alternate means, such as juxtaposition. See section 5.4.5 below for examples from the corpus illustrating this occurrence and discussion concerning its motivation. For examples of relative clauses and further discussion on the topic of noun phrase complexity, see §7.2.

5.4.4.3 Complement Clauses

Complement clauses following verbs and other constituents were grouped together as one feature for the statistical analysis. One previous study has found that complement clauses are more frequent in speech than writing (Beaman 1984). Chafe (1982) found the opposite and concluded that complement clauses provide a means of integrating information into intonational units, a tendency more common in writing than in speech.

Similarly, Ochs (1979) considers complement clauses to be a feature of planned rather than unplanned language. Greenbaum, Nelson and Weitzman (1996) found more finite complement clauses in speech but that non-finite complements were more common in writing. Biber (1998 *inter alia*) treats complementation as a way of elaborating information, and found them often implicated in expressions of personal stance. He reports that complement clauses appear most frequently in registers that are informational but produced on-line, such as interviews and speeches.

Data. There was a significant difference between the registers for complement clauses ($F = 8.133, p < .001$). The data is summarised in Table 25.

Table 25: Results: complement clauses

Register	Mean
Conversation	9.87
Radio Interview	14.00
Sports Reportage	10.92
Traditional Narrative	12.26
SPEECH	11.64
Fiction	9.41
Formal Prose	10.53
News Scripts	23.25
Popular Writing	16.80
WRITING	15.46
Total	13.55

Significant pairwise comparisons obtained only between news scripts, which had the highest occurrence level, and all of the other registers apart from popular writing and radio interview:

Table 26: Sorted data: complement clauses

23.25	●	News Scripts
16.80	●	Popular Writing
14.00	●	Radio Interview
12.26	●	Traditional Narrative
10.92	●	Sports Reportage
10.53	●	Formal Prose
9.87	●	Conversation
9.41	●	Fiction

Commentary. It seems, in Gaelic, that there is only a slight difference between the modes for complement clauses: as Chafe found, there were more complement clauses in writing. Biber's results are marginally relevant to this data; news scripts are informational and produced with little chance for editing, but they are not at all involved with the marking of personal stance, except where the speech of others has been reported. The lowest occurrence level was found in fiction, which is non-informational and highly edited. It may be that having an information orientation is more of a determiner here than planning time. Additional data seems necessary to either confirm or refute the findings of previous researchers as relevant for Gaelic. Overall, occurrence of complement clauses is not a powerful indicator of register differences in the corpus, except insofar as it helps to distinguish news scripts.

Biber *et al.* (1999) state that *that* complement clauses are commonly used to relate a person's speech, beliefs, opinions, or feelings. This kind of reporting features prominently in news scripts as the following two examples illustrate. The heads initiating the complement clauses have been placed in bold type while the complements themselves are in brackets:

- (32) *Thuir* Rùnaire nan Dùthchannan Cèin – Robin Cook – [gun robh e airson a **dhèanamh follaiseach** [nach robh e idir **fior** [gu robh Breatainn coma mu na sgìrean sin]]].
'The Secretary of State for Foreign Affairs – Robin Cook – said [that he wanted to **make it clear** [that it was not **true** at all [that Britain was unconcerned about those locales]]].'
(33) *Tha dùil* [gun do chaochail còrr is trì mìle neach] ...
'It is **expected** [that more than three thousand people have died] ...'

Example (32) begins with the past tense of the verb 'say' *thuir*, which is the most common complement-initiating element in the news scripts. Within the complement clause that follows, the verb complex *dhèanamh follaiseach* initiates another complement clause, in which the adjective *fior* initiates yet another complement clause. Biber *et al.* (1999) say that such complement clause series are not unusual in non-fiction writing. Example (33) contains the most frequent non-verbal element for initiating complement clauses in Gaelic news scripts, the noun *dùil* 'expectation'. So, much of the impetus for using complement clauses in this register is bound up with reporting speech and discussing probabilities.

5.4.4.4 Concessive Clauses

In ScG, concessive clauses are introduced by the subordinator *ged* 'although'. Previous research has indicated that concessive clauses are more frequently found in written registers (see Biber 1988) and that they tend to be associated with persuasive discourse.

Data and commentary. The Kruskal-Wallis test showed a significant difference between the registers for concessive clauses ($X^2 = 23.688, p = .001$). There was also a significant difference between the modes ($X^2 = 8.476, p = .004$). Overall, it is a rare feature, as can be seen in the generally low occurrence levels below.

Table 27: Results: concessive clauses

Register	Mean
Conversation	.34
Radio Interview	.40
Sports Reportage	.20
Traditional Narrative	.47
SPEECH	.35
Fiction	2.02
Formal Prose	1.72
News Scripts	.31
Popular Writing	.75
WRITING	1.13
Total	.74

The results generally concur with Biber (1988), which indicated a higher prevalence of concession in writing *versus* speech. However, the relationship between persuasive

registers and this feature can not be explored on the basis of the current data. Although fiction had the highest level, it is not a particularly persuasive register. Additional data is needed to increase the statistical power to the point that pairwise comparisons, not to mention parametric analysis, would be viable before making additional conclusions.

5.4.4.5 Conditional Clauses

Previous research has indicated that conditional subordination tends to positively correlate with persuasion oriented registers (Biber 1988; Jang 1998). Biber (1988) states that several researchers (e.g. Beaman 1984; Ford and Thomson 1986) have observed a higher level of conditional clauses in speech than writing but that the functional reasons for this were unclear. Ford and Thomson (1986) indicate that conditional clauses can have different functions when they are proposed *versus* postposed.

Data and commentary. A significant difference was found for conditional clauses in the registers ($F = 3.445$, $p < .003$). The only significant pairwise comparison, however, was between fiction (4.20) and sports reportage (.68). Additionally, there was not a large difference between the modes, as indicated in past research (Beaman 1984; Ford and Thomson 1986).

Table 28: Results: conditional clauses

Register	Mean
Conversation	2.29
Radio Interview	3.77
Sports Reportage	.68
Traditional Narrative	2.45
SPEECH	2.22
Fiction	4.20
Formal Prose	1.54
News Scripts	1.08
Popular Writing	2.60
WRITING	2.28
Total	2.25

It is unclear why fiction would have the highest level of conditional clauses in the eight registers when it would seem to not have an overtly persuasive function. Based upon this data, conditional clauses are not a reliable indicator of register variation and will not be discussed further in any depth.

5.4.4.6 Cause and Reason Clauses

In ScG, there are a large number of subordinators available for expressing cause and reason (see Appendix 1 and below). Jang (1998) found that causative subordination was a characteristic of more involved, exposition-based registers. Biber (1988) also found that it occurred more often in involved styles. Beaman (1984) found more causative subordination in speech, perhaps also due to its generally higher level of involvement. Miller and Weinert (1998) discovered that proposed reason clauses tend to have a more rigid word order than postposed ones.

Data and commentary. In the current data, there was no significant difference for cause and reason clauses when the grouping variable was register ($X^2 = 2.293$, $p = .109$ n.s.). Neither was there a significant difference between the modes ($X^2 = 13.869$, $p = .054$ n.s.). Table 29 presents the descriptive statistics for this feature:

Table 29: Results: cause and reason clauses

Register	Mean
Conversation	.56
Radio Interview	1.85
Sports Reportage	.30
Traditional Narrative	.96
SPEECH	.87
Fiction	1.50
Formal Prose	1.15
News Scripts	.70
Popular Writing	1.62
WRITING	1.24
Total	1.05

Although the higher level of cause and reason clauses in radio interview points to the kind of interpretation offered by Biber and Jang above, additional data, for the purposes of increasing statistical power, is needed before we can be more conclusive about their distribution in Scottish Gaelic registers. Possibly, one finds more of this feature in radio interview due to a higher level of explanations and descriptions in response to the interviewer's questions. However, it is not a robust predictor of variation in this data.

The most common subordinators used for expressing cause and reason overall were *oir* and *a chionn* ('s). These were overly prominent in the written registers. Other subordinators found in the corpus included the following, which were fairly evenly distributed in the spoken texts: *air sgàth*; (*air*) *sàilleadh*; (*bh*)*on*; *a thoireadh*/ *a thoradh*; *air dhòigh*; *airson* ('s); *leis*; *ri linn* ('s); *seach*.

5.4.5 Juxtaposition versus Subordination in Spoken Scottish Gaelic

The results of the previous section on subordination indicated a generally greater prevalence of subordinate clauses in writing than in speech. This was especially the case for relative clauses, concessive clauses, and, to an extent, complement clauses. The current section explores one possible explanation for this finding: juxtaposition as used in speech can fulfil some of the same functions covered by subordination in writing.

Spoken language is able to be less lexically explicit in how it handles clausal relations. While speakers have pitch, amplitude and speed in their repertoire to foreground or background clauses in relation to each other, writers must resort to lexico-grammatical strategies to accomplish the same functions. In spontaneous spoken ScG, we see cases of two juxtaposed independent clauses with one of them clearly more prominent or thematic and the other, functioning to provide background information. Three examples are offered below of juxtaposed clauses in speech. In each, the second clause would probably have been cast as subordinate had the utterances come from written rather than spoken language.

- (34) agus chì thu ri taobh an taigh' aca ars' esan air a' mheadhan-latha
and sec-INDEX 2S beside ART house-G at-3P said 3S-M on ART middle-day-D
'and you can see beside their house he said in the middle of the day'

no air a' mheadhan-oidhche tha an stòbh aca geal
or on ART middle-night be-PRES ART stove-N at-3P white
or middle of the night their stove is white hot [Lit. 'white']

- (35) tha mi a' creidsinn bha e uairigin aig Baile Raghnaill e fhèin
be-PRES 1S PROG believing-VN be-PAST 3S-M once at Town Ronald-G 3S-M REFL
'I am believing he was once at Balranald himself'

- (36) chanainn an duine aig a' chùl tha e cianail important dhan team
 say-INDEF2-1S ART man at ART back be-PRES 3S-M awfully important to-ART team
 'I would say the person at the back he is awfully important to the team'

In example (34), the first clause is perceived as incomplete: in information-processing terms the addressee is told to 'open a file' for 'what the addresser could see beside their house...'. In the second clause, the information necessary to fill in the file opened in clause 1 is provided – 'their stove is white' – but this is expressed as an independent clause. In writing, it would have likely been presented as a complement: *chì thu ... gu bheil an stòbh aca geal* 'you can see... that their stove is white hot'. The verb in the written version (*bheil*) is the dependent form of the verb 'to be'. The subordinate relationship is lexicalised through this verb form and the complementiser *gu(n)* which precedes it. There is nothing grammatically 'wrong' with using an independent clause here; the relation holding between the clauses is less explicit but the addressee has no problems in perceiving it.

Similarly, in example (35), the first clause is also felt to be incomplete. When one hears 'I am believing', one expects a grammatical object; *creid* 'believe' is not an intransitive verb. The second clause provides this object, but it is not morphologically coded as a complement. Again, if it had been a written utterance, it probably would have obtained as *tha mi a' creidsinn gu robh e aig Baile Raghnaill e fhèin* 'I believe that he was at Balranald himself'. Once again, the new independent clause, instead of prompting the listener to 'open another file', provides the missing piece as the previous clause is interpreted as incomplete. Example (36) begins with the verb (with pronoun suffix) *chanainn* 'I would say', which normally takes a complement clause. However, the speaker avoids this construction and instead uses an extracausal NP for the purposes of reference ('*an duine*'), followed by a clause with a pronominal subject that is coreferential to the previous NP. This is classic left-detachment. The important aspect here is that the speaker does not enlist a complement clause, but instead, a main clause. In effect, this clause provides the object of *chanainn*, but is expressed as morphosyntactically non-integrated.

In the examples above, spontaneous speech was seen to 'avoid' complement clauses. All three provide a clausal object, not in a subordinate form, but rather, as a main clause. Juxtaposition of clauses and phrases can also function similar to relative clauses, providing additional information about a noun phrase. Consider example (37), which also illustrates the tendency in spontaneous speech for long stretches of syntactic material to sometimes intervene between semantically related clauses or other constituents:

- (37) ¹*chuala mi* ²*am boireannach* ³*bha mi ag ràdh riut an-dè* ^(2 cont.) ⁴*a chunna mi an Griminis an-shiudach* ⁵*bha i sa bhunkhouse againn* ⁶*thall an-shiudach chunnaic mi i* ⁷*tha an càr aice an Griminis fhathast*
 'I heard ²the woman ³I was saying to you yesterday ^(2 cont.) that I saw in Griminish there ⁴she was in our bunkhouse ⁵over there I saw her ⁶her car is in Griminish still

This example begins with the words *chuala mi* 'I heard' but the grammatical object is not offered. Before we learn what the speaker heard – provided in clause 6 (*tha an càr aice an Griminis fhathast*) – he first spends four clauses to build up information regarding the woman who owns the car. This is a good example of how spontaneous spoken language builds up descriptions gradually, in layers. (These two parts, clause 1 and clause 6, are in bold type to indicate their relationship to one another.) There is only one relative clause – clause 2, beginning with *am boireannach* – and it is cleaved by the independent clause *bha mi ag ràdh riut an-dè*. The two underlined parts above are obviously related to each other;

am boireannach is the NP head and *a chunna mi...* is the modifying relative clause. The new, important information to which the speaker is ultimately leading – that the woman's car is still in Griminish – only occurs in clause 6 once the addresser is certain that the addressee knows who he is speaking about. As in examples (34)–(36) above, this clause provides the information needed to fill in the file opened by the first clause, but it is in the form of an independent clause rather than a complement clause. Moreover, although four clauses serve to provide background information about *am boireannach* 'the woman', only one relative clause is used. Despite the fact that few of these clauses evince markers of any kind relating them back to this noun phrase, their elaborating function is still clear.

What juxtaposition loses in explicitness, it gains in flexibility. It would be possible to recast the above as more syntactically integrated, incorporating relative clauses, but a linearly dependent syntax introduces parsing difficulties in this case. Consider example (38) below:

- (38) *'chuala mi 'gu bheil an càr aig a' bhoireannach 'a bha mi ag ràdh riut an dè 'a chunnaic mi ann an Griminis ann a-shiudach 'a bha sa bhunkhouse againn ann an Griminish fhathast*
 'I heard 'that the car of the woman 'who I was saying to you yesterday 'who I saw in Griminish there 'who was in our bunkhouse is in Griminish still

Most of the syntagms occur in the same order as example (37), but to force syntactic completeness, clause 6 of example (37) above (clause 2 here) is integrated as a complement of the first clause. (Readers may also notice that clause 5 of the prior example has been omitted, for the reason that the information is redundant.) Without using an extracausal NP, to separate role from reference, the resulting object NP of the preposition *aig* becomes very dense, introducing an unacceptable number of intermediary elements before the predating phrase *ann an Griminis fhathast* 'in Griminish still'. So, while juxtaposition is syntactically inexplicit, it enables the incorporation of more information about a referent than is possible with a series of relative clauses.

In order to recast this utterance in more polished prose, it is necessary to significantly alter both its syntax and information structure:

- (39) *'Bha mi ag ràdh riut an-dè 'gum faca mi am boireannach 'a bha sa bhunkhouse againn ann an Griminis. 'Chuala mi 'gu bheil an càr aice ann an Griminis fhathast.*
 'I was saying to you yesterday 'that I saw the woman 'who was in our bunkhouse in Griminish. 'I heard 'that her car is Griminish still.

The most significant features of example (39) are its two-sentence structure and its greater number of dependent clauses. In example (37) there is only one dependent clause – a relative clause – whereas there are three in example (39): two complement clauses and a relative clause. Although these dependent clauses manifest greater semantic explicitness, they change the information structure of the original message. Example (39) presents *bha mi ag ràdh riut an-dè* 'I was saying to you yesterday' in the first position, thus giving it more prominence than example (37) did, in which it was only a parenthetical comment. Where (37) uses a left-detachment construction (*am boireannach ... bha i sa bhunkhouse againn*), example (39) has integrated the NP as the object of a complement clause (*gum faca mi am boireannach*). Furthermore, while (39) uses a relative clause to express the fact that the woman had been in their bunkhouse, example (37) uses an independent clause, slotted on to the previous ones. Finally, example (37) presents the new, important information, that to which the whole utterance is leading – the woman's car is still in Griminish – as an independent clause. This clause stands in bold contrast to the rest of the utterance: the listener is waiting for the speaker to provide the complement to *chuala mi* and here it is at last. Example (39) lacks the build up present in example (37), also this

information is provided as a complement clause rather than an independent clause, diminishing its impact. Overall, the written version is more logically expressed; constituents are clearly marked as relating to one another. However, its information structure lacks some of the expressiveness of example (37).

To conclude, juxtaposed clauses in speech can often fulfil functions usually reserved for subordinate constructions in writing. This characteristic of spoken language may, in part, explain the lower amount of subordination generally found in spoken registers, as indicated in the current study. What are formed normally as morphosyntactically overt complement clauses in writing are often in the form of juxtaposed independent clauses in speech. Moreover, juxtaposition is capable of providing additional clause-based information about an NP referent, a function usually reserved for relative clauses in writing. Both uses of juxtaposition are helpful for the building up of information gradually in speech, to assist with human processing limitations. In this way, they can be similar to or even implicated in left-detachment constructions, as seen in (36) above. The advantage of using juxtaposition is greater syntactic flexibility, but explicitness is compromised, which would be unacceptable in most forms of written language.

5.5 Summary of Chapter Five

This chapter surveyed the distribution of a number of features related to information structure and clause type variation in Scottish Gaelic. Beginning with the information structure features, no solid findings were obtained for *'s e* and *'s ann* clefts, in contrast to previous research on IT clauses in English. However, **demonstrative clefts** were more common in speech, especially in conversation, while they were almost absent from news scripts. Due to this distribution, it was suggested that demonstrative clefts are a sensitive marker of the level of discourse freedom in a register. The proportion of **left-detachment** in the registers is clearly weighted towards spontaneous speech. In light of this result and previous research, it appears that left-detachment constructions are a product of the more stringent production constraints of spontaneous speech. They may also be indicative of an attempt to aid addressee comprehension through the separation of role and reference. **Right-detachment**, which was generally less common than left-detachment, was most common in the conversation register, but absent from radio interview. However, it appeared with an even greater frequency in fiction than did left-detachment. More research is needed into the discourse motivations of right-detachment in Scottish Gaelic. **Adverbial fronting** was most common in the written registers, especially those which tended to be discursive (e.g. formal prose) rather than reportage-based (e.g. news scripts). In terms of **subordinate clause placement**, most registers preferred postposed placement of subordinate clauses (after the matrix clause). This was particularly marked in the news register, where only 7% of subordinate clauses occurred before the matrix clause. It was suggested that this finding reflected news scripts' strong forward inertia and discontinuous nature; there was less need for devices that guidepost transitions between clauses.

A number of features associated with clause types were found to indicate robust register differences. **Clausal coordination** was found to be most prominent in spoken monologue, such as traditional narrative and sports reporting. In these registers, it was suggested that coordination was used to maintain the pace of the narrative and, possibly, to aid information processing by dividing up individual clauses and phrases via explicit coordinators. Although the discourse functions of **clausal cosubordination** are unknown at present, it was clearly a marker of narrative text, as well as being common in the news scripts. Further research is necessary before any claims can be made concerning its motivation in such texts. The proportion of **main/matrix clauses** was highest in conversation and narrative (especially traditional narrative) and lowest in formal prose and news scripts. It was maintained that a higher proportion of this feature was linked with smaller, simpler clauses and fewer clause combinations, both of which have been found to be associated with more typically oral texts.

The findings for subordinate clauses found a wide range of variation in the distribution of features which are sometimes lumped together in other studies. **Temporal subordination** was most prominent in narrative, where time is dynamic and relative, built up and stabilised through chains of clauses occurring in relation to one another. As found in prior studies, **relative clauses** were most frequent in writing, especially formal prose and popular writing, and less common in the spontaneous spoken registers. This distribution indicates that relative clauses are one of the features sensitive to production constraints. **Complement clauses** were generally more frequent in writing than speech, and especially common in news scripts. Although this feature is not a particularly robust marker of Scottish Gaelic register differences, it seems to indicate, as found in other studies (e.g. Biber 1998), an informational orientation.

Finally, asyndetic clausal juxtaposition was surveyed. It was concluded that this type of construction in speech can fulfil some of the same functions reserved for clausal subordination in writing. Importantly, it is also able to provide syntactic flexibility in the face of the greater processing limits of spontaneous speech.

6 Morphosyntax and the Lexicon

As opposed to the previous chapter, which examined different clause types and aspects of information structure, the current chapter concentrates upon features within the clause. The features represented here are an assorted group, but all have been targeted in past research on other languages as being relevant to register variation, or included as a result of observations that were made during an initial survey of the Scottish Gaelic texts. As in the previous chapter, each feature will be introduced with reference to previous research, followed by the data, examples where relevant, and interpretation.

6.1 Verbal Morphosyntax

This section concentrates on the tense-aspect-modality system and valence and how they vary throughout the registers. The features surveyed primarily involve the way that time, speaker orientation to propositions, and the presentation of arguments are encoded in the morphosyntax of the language. The selection of these features has a great impact on the surface ordering of elements in the Scottish Gaelic clause. For examples of these effects, refer to A§2.3.1.2 (tense-aspect-modality system) and A§2.3.2.2 (valence decreasing expressions) in the grammar.

6.1.1 Aspect

Gaelic aspectual marking falls into three main types: perfect, progressive, and prospective. The first two are clearly prevalent in the data but prospective marking is somewhat rare. Table 30 presents the means for these three features:

Table 30: Results: aspect

Register	Perfect	Progressive	Prospective
Conversation	1.27	24.92	.52
Radio Interview	3.17	32.13	.69
Sports Reportage	3.65	42.46	.10
Traditional Narrative	1.66	19.63	.47
SPEECH	2.40	29.66	.43
Fiction	3.81	23.10	.61
Formal Prose	4.97	16.46	.81
News Scripts	8.13	25.14	.56
Popular Writing	6.32	23.01	.91
WRITING	5.95	22.04	.73
Total	4.17	25.85	.58

Perfect aspect expresses a state in the 'present' that has come about from an earlier situation, usually an activity, as coded by the verb. In other words, perfect aspect reaches back to something that happened in the past which is still relevant within the temporal frame of a given proposition. Scottish Gaelic codes perfect aspect usually by using the preposition/aspect marker *air* 'after' before a verbal noun, as in (40):

- (40) *Ach chan eil mòran earbsa air a bhith agam bho chionn iomadh latha ann an nì a their poileatiseans.* (POP)
 But I haven't had much trust since many a day in a thing that the politicians say [Lit. 'Not much trust has been after being at me ...']

Perfect aspect has been researched by Biber (1988) and Jang (1998) in terms of register variation. Biber found that it was more prevalent in narrative registers while Jang found that interactional registers had the highest incidence.

Data. A significant difference obtained between the registers for perfect aspect ($F = 7.542$, $p < .001$). There were four significant pairwise comparisons as illustrated below:

Table 31: Sorted data: perfect aspect

8.13	●	News Scripts
6.32	●	Popular Writing
4.97	●	Formal Prose
3.81	●	Fiction
3.65	●	Sports Reportage
3.17	●	Radio Interview
1.66	●	Traditional Narrative
1.27	●	Conversation

The arrows indicate contrasts between news scripts and popular writing on the upper end of the data and traditional narrative and conversation on the lower one. There was a wide range evinced, with news scripts having nearly seven times the amount of perfect aspect as conversation.

Commentary. This result is surprising considering the low levels evinced for traditional narrative and conversation, which are the most overtly narrational and interactional of the registers respectively. In effect, the findings are diametrically opposed to those of Jang and Biber. It is possible that news scripts may evince a high level of perfect aspect due to the influence of English. In translation from English, which encodes a more distinct contrast between perfect aspect and the simple past, it seems to occur with greater frequency than it would naturally: Perfect aspect can be conveyed in Gaelic simply by the use of the past-definite tense. On the other hand, news scripts may have the highest return for perfect aspect because they commonly deal with occurrences that began in the past but are still relevant in the present. Overall, the above findings indicate that perfect aspect is more common in written registers. Perhaps, there is also a connection, as Payne (1997: 173) mentions for other languages, between perfect aspect and passive voice in Scottish Gaelic (for an example see (231)).

Progressive aspect indicates a mutable, on-going process. In ScG, it is coded by the presence of *ag* 'at' before a verbal noun. Previous research (Biber *et al.* 1999) has found progressive aspect to be most common in conversation and least common in academic writing.

Data and commentary. There was a significant difference between the registers ($F = 10.01$, $p < .001$). The most robust pairwise comparisons obtained between sports reportage and radio interview on the higher end of the scale and formal prose and traditional narrative on the lower end. The large range indicates robust differences between the outlying registers.

Table 32: Sorted data: progressive aspect

42.46	●	Sports Reportage
32.13	●	Radio Interview
25.14	●	News Scripts
24.92	●	Conversation
23.10	●	Fiction
23.01	●	Popular Writing
19.63	●	Traditional Narrative
16.46	●	Formal Prose

This result points to high levels of progressive aspect being characteristic of registers that are concerned with on-going events. Sports reportage is perhaps the best characterisation of this kind of register in the corpus, being an on-line description of spontaneously occurring events. The lowest levels are found in formal prose, which is perhaps the most abstract of the registers, and not at all bound up with current reference. A comparison of these two registers shows the difference in aspectual assignment:

Example:

Gough a' sealltainn cò th' aigesan Gough a' feuchainn tè àrd suas gu fad 's a bheil Johnson ach tha Yositch ann an siud còmhla ris agus choinnich Johnson i 's chuir e a-mach i agus 's e sàdadh a-steach a bhios ann do dh'Alba dìreach lethach slighe a-steach air taobh leth Yugoslavia Alba a' cluich bhon a' làmh chli chun a' làimh dheis mar a tha sinn a' coimhead air agus 's e Nicol fhèin a tha 'dol dha gabhail Nicol leatha [name] a' ruith air a shon tha Johnson Johnson dìreach air oir a' bhucais ga cur air ais gu Nicol (SPORTS: 101 words)

Translation:

Gough **looking** who he has Gough **trying** a high one up to where Johnson is but Yositch is there with him and Johnson met it [the ball] and he put it out and it's a throw in that will be there for Scotland exactly half way in on Yugoslavia's half Scotland **playing** from the left hand to the right hand as we are **looking** on it and it's Nicholl himself who is **going** to take it Nicholl with it [name] **running** up for it but Johnson Johnson just at the corner of the box **putting** it back to Nicholl

Example:

Tha na sgoilearan dhen bhàrail gur h-ann man tritheamh ceud deug a thòisich an cànan Èireannach air dol na dhualchainntean agus gu freagradh sin air Gàidhlig Alba cuideachd. Ach gu b' e gu dè theirear mu Èirinn is gann a ghabhas e chreidsinn nach robh tàrmachadh cainnte cinneachadh fada roimhe sin an corra chèarna de Alba am measg muinntir na tuatha – ga nach biodh na adhbhar air sin ach gur h-e coigrich dhan Ghàidhlig air thùs a bh' ann am mòran dhiubh. Bhiodh smachd foghlaim air filidhean agus air pearsachan eaglais gan cumail-san air an aon ràmh, ach is cinnteach gum biodh saorsa aig a' chòrr dhen t-sluagh mar a gheibhear air feadh an domhain. Co-dhiù, sa bhliadhna 1258, nuair a bhruidhinn Domhnall Og 'tria san nGaoidhilc nAlbanaigh', tha e soilleir gu robh roinn eadar cainnt Alba agus cainnt Èirinn da-rìribh. Ach chan eil iomradh sam bith ann gu robh e doirbh dhaibh a chèile thuigsinn. (FORM: 155 words)

Translation:

The scholars are of the opinion that it is about the thirteenth century that the Irish language began to break up into dialects and that that is appropriate as well for the Gaelic of Scotland. But whatever is said about Ireland, it can hardly be believed that dialect divisions were not **growing** far before that in the odd Scottish locale amongst the common folk – even if there were no other reason for this than the fact that many of them were strangers to the Gaelic. The effects of education on the educated elite and the heads of the church would be **keeping** them along the same [linguistic] path, but it is certain that the rest of the populous would have had freedom, as is found universally. Anyway, in the year 1258, when Young Donald spoke 'through the Scottish Gaelic', it is clear that there was a division between the speech of Scotland and the speech of Ireland indeed. But there is no indication that it was difficult for them to understand each other.

In the example, from sports reportage, there are seven instances of progressive aspect, all but one of which (*'dol dha gabhail* 'going to take it', which is an instance of the prospective) refer to spontaneously occurring actions in linear progression. Events on-field are continuous, each discrete event rolling into the next one. Progressive aspect is necessary to accurately describe such a situation. In contrast, in the formal prose example, there are only two cases of progressive aspect despite there being 1.5 times the number of words. This indicates that this register is not concerned with on-going events, but rather, with much more static subject matter.

Prospective aspect is used to indicate an event that is about to happen. It can be coded through various means in ScG, usually by employing *a' dol* 'going' + infinitive, *gus* 'to' + infinitive or *an impis* 'almost' or 'about to' + verbal noun. This feature has apparently not been implicated in previous register variation studies.

Data and commentary. Differences in levels of prospective aspect were tested with the Kruskal-Wallis statistic and found to be non-significant ($X^2 = 7.6$, $p = .369$ n.s.). Differences in mode were also non-significant ($X^2 = 3.076$, $p = .079$ n.s.). Clearly, additional data is needed to increase the statistical power of this comparison and it will not be discussed again in this study.

6.1.2 Tense

Table 33 displays the means for the four variations of 'tense' in Scottish Gaelic (see comments in §2.3.1.2, Appendix 1): present definite; past definite; indefinite 1 ('present-future'); indefinite 2 ('past-habitual'):

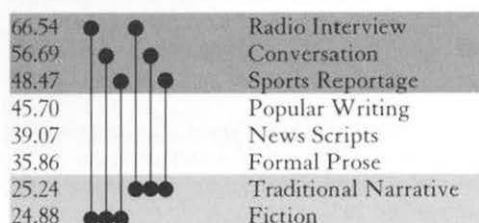
Table 33: Results: tense

Register	Present	Past	Indef1	Indef2
Conversation	56.69	49.84	13.48	11.53
Radio Interview	66.54	30.39	16.96	10.86
Sports Reportage	48.47	39.25	6.93	1.83
Traditional Narrative	25.24	82.71	14.07	15.68
SPEECH	48.32	51.61	12.64	9.93
Fiction	24.88	68.22	11.49	21.74
Formal Prose	35.86	46.04	7.54	10.26
News Scripts	39.07	34.64	14.22	5.87
Popular Writing	45.70	39.46	11.55	13.34
WRITING	37.24	45.81	11.29	12.42
Total	42.78	48.71	11.97	11.17

Biber (1988) maintains that **present tense** most often occurs as a feature of involved registers as it deals with current reference. However, he says that it can also be "used in academic styles to focus on the information being presented and remove focus from any temporal sequencing" (p. 228). Biber *et al.* (1999) found that present tense occurred most frequently in conversation, with much lower levels in fiction. They state that present tense in conversation indicates a concentration on immediate context and can also be used to denote current habitual behaviour. Although the latter function was historically carried by the first indefinite ('present-future') in Gaelic, contemporary usage favours the present, apparently through influence from English. Gaelic has only one verb capable of expressing unambiguous present tense: the verb 'to be' *tha*, which often acts as an auxiliary in combination with a verbal noun.

Data and commentary. A significant overall difference obtained for present tense ($F = 8.94$, $p < .001$). A number of significant pairwise were also found:

Table 34: Sorted data: present tense

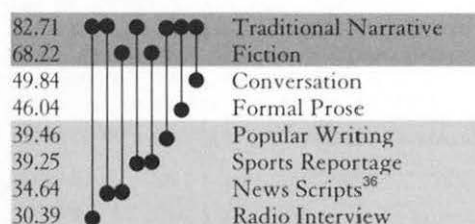


Radio interview clearly has the greatest instance of present tense followed by conversation and sports reportage, while fiction and traditional narrative have the lowest occurrence levels. The arrows indicate significant differences between these two main groups, highlighted by the shading in Table 34. There is a unequivocal division between registers concerned with immediate context and those of a more abstract or narrative nature on this feature. Notably, three of the four spoken registers assume the highest positions for occurrence of present tense.

Past tense is the classic superficial marker of narrative registers (see Biber 1988; 1995). Biber *et al.* (1999) find past tense most prevalent in fiction, with moderate amounts in news. Tannen (1980) reports one major difference between speech and writing in past reference: written narrative tends to maintain past tense marking throughout while spoken narrative often begins with past tense marking and then shifts to the 'historical present'.

Data and commentary. There was a significant difference between the registers for past tense verbs ($F = 8.964$, $p < .001$). There were also a number of significant pairwise comparisons, as presented in Table 35:

Table 35: Sorted data: past tense



Traditional narrative and fiction together have, by far, the highest occurrence level of past tense. Radio interview on the other hand, which was found to have the highest level of present tense, has a very low level of past tense. Interestingly, conversation had the third highest level of past tense out of the registers, while it also had a high level of present tense, as seen above. This seems to indicate that it is a very 'verbal' register and that there

³⁶ One would expect, perhaps, news scripts to have evinced more past tense. Two explanations may be offered for this finding: there is general tendency for clauses in news scripts to be longer (see result for main/matrix clauses §5.4.3), with less finite verbs overall compared to some other registers; tense is not as fixed in news scripts as in other kinds of narrative based registers—the news today is as much concerned with events that are still to happen or in progress as it is with events that have happened.

Passive voice in Gaelic can take various forms. These are discussed, with examples, in section 2.3.2.2 of Appendix 1. The main division usually encountered in the literature is between so-called periphrastic (or 'analytic') passives and inflected (or 'synthetic') ones. In periphrastic passives, meaning is spread out over a number of words. An example of a periphrastic passive form in Gaelic is *chaidh a thogail* 'it was raised' [Lit. 'its raising went']. In the inflected counterpart for this utterance, meaning is conveyed through a single main verb with suffix, and a pronoun: *thogadh e* 'it was raised'. Lamb (1999) discovered that the use of inflected impersonals declined in Scottish Gaelic news scripts between the 1950s and 1990s and that there was a concomitant increase during this time in morphosyntactic irregularity and progressiveness, seen especially in case marking. Whereas a 1965 script has *Dh'innseadh gum b' iad an dithis a chailleadh* a script from today would most likely have *chaidh innse gur iad an dithis a chaidh a chall*. The changes are: *dh'innseadh* (inflectional) → *chaidh innse* (periphrastic); *gum b' iad* (past copula) → *gur iad* (present copula); and *chailleadh* (inflectional) → *chaidh a chall* (periphrastic).

Agents are specified in passive constructions using, mainly, the prepositions *le* 'with' or *aig* 'at'. Traditionally, *le* was used with INSTRUMENTS and *aig*, with AGENTS. The following examples are passives with an INSTRUMENT and AGENT specified:

- (41) *chaidh a mharbhadh le gunna* 'he was murdered with a gun'
 (42) *chaidh a mharbhadh aig Teàrlach* 'he was murdered by Charles'

Data and commentary. The means from the passive data are presented in Table 37:

Table 37: Results: valence decreasing constructions

Register	+ Agent	Synthetic	Total
Conversation	.00	.21	1.12
Radio Interview	.00	1.92	4.42
Sports Reportage	.00	.20	2.13
Traditional Narrative	.29	3.17	4.14
SPEECH	.08	1.35	2.87
Fiction	.24	.57	2.70
Formal Prose	.61	6.60	11.37
News Scripts	2.13	.66	16.84
Popular Writing	.51	1.38	6.60
WRITING	.90	2.26	9.60
Total	.49	1.80	6.24

A significant difference in the registers obtained for **passive clauses with agents** ($F = 9.164$, $p < .001$) although few pairwise comparisons were significant:

Table 38: Sorted data: passive clauses with agents

2.13	●●●●	News Scripts
.61	●●●●	Formal Prose
.51	●●●●	Popular Writing
.29	●●●●	Traditional Narrative
.24	●●●●	Fiction
.00	●●●●	Conversation
.00	●●●●	Radio Interview
.00	●●●●	Sports Reportage

Passives with agents specified are rare overall but, as indicated by the arrows, news scripts diverge from the rest of the corpus. The higher occurrence of the passive in this register seems to be indicative of its generally more detached, fact-based orientation.

The following three examples from the news scripts illustrate the passive with agent construction. Examples (43) and (44) are of periphrastic passives, while example (45) is inflected:

- (43) *Thèid innse do Chaledonian Mac a' Bhruthain le riochdairean Chomhairle nan Eilean.* [sic]
'It will be told to Caledonian MacBrayne by the representatives of the Western Isles Council'
- (44) *Chaidh an cù – Japanese Aketa – a chur gu bàs le lighiche-sprèidh.*
'The dog – a Japanese Aketa – was put to death by a veterinarian.'
- (45) *Tha eòlaichean a dh'fàsadh leis an riaghaltas ...*
'Experts who were employed by the government ...'

There was also a significant difference between the registers in terms of **total passive clauses** ($F = 25.55, p < .001$). Post-hoc analyses revealed the following significant pairwise comparisons:

Table 39: Sorted data: total passive clauses

16.84	●	News Scripts
11.37	●	Formal Prose
6.60	●	Popular Writing
4.42	●	Radio Interview
4.14	●	Traditional Narrative
2.70	●	Fiction
2.13	●	Sports Reportage
1.12	●	Conversation

Once again, the news scripts have the highest level of the feature, followed by formal prose, while conversation has the lowest occurrence level. It is notable that three of the top four register are written ones: news scripts; formal prose; and popular writing. Fiction stands in contrast to these three registers.

The arrows indicate that news scripts and formal prose as a group are significantly different from most of the other registers. A higher occurrence of passives in Gaelic apparently distinguishes more informational and detached texts such as news scripts and formal prose from more involved ones. This result is similar to that obtained by Biber (1988) and Chafe (1982).

Biber *et al.* (1999) state that the occurrence of passives in news scripts is partly conditioned by the subject matter: AGENTS in news scripts are often able to be recovered through cultural frameworks – knowledge to do with courts of law or hospitals, for example. They are also often recoverable from surrounding discourse (as in example 1 below). In academic registers, on the other hand, the suppression of AGENTS is due primarily to the fact that the register is detached, being more concerned with providing generalisations than naming the specific person effecting an action. Examples illustrate these factors in the Gaelic texts. Passive/impersonal verbs have been put into bold type and AGENTS recoverable from the text have been underlined.

- (46) News script 1
Chaidh leudachadh a-rithist a dheanamh air a' chasg air iasgach chreachain air taobh siar na h-Alba air sgàth 's a' puinnsean ASP. Tha riaghaltas na h-Alba air an casg a sgaoileadh a-mach gu deas air Diuraidh agus Ile.

'An expansion **has been done** again on a ban on shellfish fishing on the west coast of Scotland because of ASP poisoning. The **Scottish government** has widened the ban out to the south of Jura and Islay.'

(47) News script 2

Chaidh casaid muirt a chur air Buford Furrow – an duine a bha ga cheasnachadh mu ionnsaigh air ionad Iubhach ann an Los Angeles. Chaidh fear a' phuist a mhurt goirid as dèidh na h-ionnsaigh a's an deach còignear a leòn.

'A murder charge **has been put** on Buford Furrow – a man who was questioned about an attack on a Jewish centre in Los Angeles. A postman **had been murdered** shortly after the attack in which five people were wounded.'

(48) Formal prose 1

Tha ciadad àrd de na Sinich aig a bheil an seòrsa fala B, seòrsa nach fhaighear a-measg nan Innseanach an Ceann-a-tuath 's an Ceann-a-deas Ameireagaidh.

'There is a high percentage of Chinese who have the blood type B, a type which is **not found** amongst the Indians in North and South America.'

(49) Formal prose. 2

A-réir coltais, chaidh an sgeulachd seo a dhealbh an toiseach an Eirinn anns an t-siathamh ceud deug.

'Apparently, this story **was created** first in Ireland in the sixteenth century.'

In example (46) above, from the news scripts, the argument left out of the first (passive) clause is able to be recovered in the second clause: *riaghaltas na h-Alba*. In example (47), the first clause simply says that Buford Furrow has been charged with murder, but does not say who made the charge. This is because this information is superfluous: nearly everyone knows that charges are made by judges and, unless the particular judge becomes important in some way, or the case becomes as infamous as that of O.J. Simpson, the public does not care who it is. The second clause in (47) says that a postman has been murdered; the implication is that it is Furrow who was the agent of his murdering.

Example (48), from a biology textbook, presents a generalisation: there is a blood type which cannot be found amongst a certain population. For the purposes herein, there is no reason to supply the person(s) who made this finding; it would be digressive and probably lead into a discussion of blood typing methodology, etc. Example (49) again presents a generalisation, but this time, it is not possible to know who the actual individual was who created the story – only that it probably happened in Ireland.

All of these examples focus on information first: that an expansion has been done; that a man was charged with murder; that a blood type cannot be found; and that a story was created. In formal prose and news scripts, information gets primary place. In other registers, AGENTS are more important and given priority over their actions.

For **synthetic passives and impersonals**, a significant difference was also found ($X^2 = 30.449$, $p < .001$). However, no significant difference obtained between the modes ($X^2 = 1.311$, $p = .252$ n.s.). Because the contrast was tested with a non-parametric measure (Kruskal-Wallis), pairwise comparisons were not possible. However, there seems to be a indisputable contrast between the outliers, as seen in Table 40:

Table 40: Sorted data: synthetic passives and impersonals

6.60	Formal Prose
3.17	Traditional Narrative
1.92	Radio Interview
1.38	Popular Writing
.66	News Scripts
.57	Fiction
.21	Conversation
.20	Sports Reportage

Interestingly, formal prose and traditional narrative have the highest occurrence levels of this feature. This finding, along with those concerning case marking (see §7.1), gives the impression that these registers are the most morphologically conservative in the corpus. Conversation and sports reportage again have the lowest occurrence here. This is not surprising considering their informality and their greater overall 'involved' orientation. Interestingly, when this result is placed beside the two others in this section, it appears that news scripts have a preference towards more analytic means of expressing passive voice, as detected in Lamb (1999). In fact, while 77% of passive constructions in formal prose are synthetic, this is true for only 6% of those in the news scripts. See Appendix 1 for examples of the different passive and impersonal constructions in ScG.

6.1.4 Clausal Negation

Clausal (or 'analytic') negation functions to reject a proposition. Biber (1988) finds that clausal negation tends to be associated with more involved, primarily oral registers as opposed to more informational ones (his 'Dimension 1'). He interprets it as a fragmented feature, characteristic of spontaneous speech. Biber *et al.* (1999) find that clausal negation occurs most frequently in conversation, a factor of the overall greater number of verbs in this register as opposed to most others.

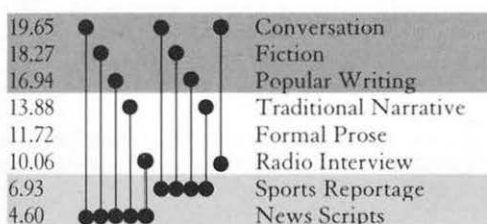
Data and commentary. A significant difference was found for clausal negation ($F = 12.337$; $p < .001$) and the means are detailed in Table 41 below:

Table 41: Results: clausal negation

Register	Mean
Conversation	19.65
Radio Interview	10.06
Sports Reportage	6.93
Traditional Narrative	13.88
SPEECH	12.76
Fiction	18.27
Formal Prose	11.72
News Scripts	4.60
Popular Writing	16.94
WRITING	12.73
Total	12.75

Several pairwise comparisons were also significant:

Table 42: Sorted data: clausal negation



Conversation had the greatest occurrence level of clausal negation followed closely by fiction. On the other hand, sports reportage and news scripts had relatively few (4.6 per 1000) instances of the feature. There were significant pairwise contrasts between a number of the registers, with particularly robust differences between sports reportage with news

scripts and nearly all of the others. The shading illustrates that the two most opposed groups are conversation, fiction, and popular writing *versus* sports reportage and news scripts.

While these results do not completely agree with Biber (1988), they are almost in parallel with Biber *et al.* (1999). Conversation evinces the highest levels of clausal negation and it is closely followed by fiction. Fiction cannot be characterised as primarily oral or fragmented, but was found in Biber (1999) to also have a high incidence of clausal negation. At the other end of the distribution are news scripts and sports reportage. Perhaps these two have low returns because they are mainly concerned with affirming that something *did* happen or *is* happening rather than the converse. The most consistent functional grouping of registers on this feature concerns the opposition between more factually-oriented registers and more imaginative or involved ones.

6.1.5 Mode: Interrogative clauses and Imperative Verbs

Both interrogative clauses and imperative verbs are superficial signs of spoken, interactive registers (Biber 1988; 1995; Jang 1998). In Scottish Gaelic, interrogative clauses are indicated by the preverbal particles *a(n/m)*, for positive questions, or *nach*, for negative questions, followed by dependent verbal morphology. Imperative verbs are formed through the use of the verbal root, plus pronominal suffix for some categories of person. See §2.3.1 in Appendix 1 for more information on these two topics.

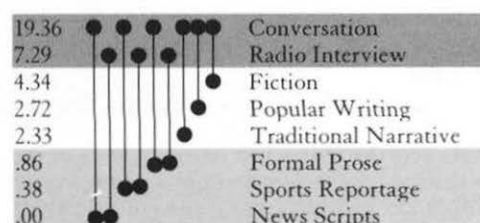
Data and commentary. There was a significant differences between the registers for interrogative clauses ($F = 22.704, p < .001$) and for imperative verbs ($F = 14.008, p < .001$). The means are detailed in Table 43:

Table 43: Results: interrogatives and imperatives

<i>Register</i>	<i>Interrogative clauses</i>	<i>Imperative Verbs</i>
Conversation	19.36	1.30
Radio Interview	7.29	.58
Sports Reportage	.38	.00
Traditional Narrative	2.33	4.49
SPEECH	7.34	1.65
Fiction	4.34	2.39
Formal Prose	.86	.22
News Scripts	.00	.00
Popular Writing	2.72	.54
WRITING	1.90	.71
Total	4.62	1.18

Pairwise comparisons of **interrogative clauses** showed several significant differences:

Table 44: Sorted data: interrogative clauses



Conversation has the highest instance of interrogative clauses by far in the corpus, followed by radio interview. At the other end of the spectrum, news scripts returned no cases of the feature. The results indicate clearly that more interactive registers have a much higher occurrence level of questions. The main contrasts between the registers, illustrated by the arrows and shading, were between conversation and radio interview and all other registers. In conversation and radio interview, most of the interrogative clauses were in the form of direct questions, from one participant to another. In popular writing, there was a relatively high number of interrogative clauses for a written register. Despite being non-interactive, popular writing is more involved and informal than registers like formal prose and news scripts. Most of the instances of interrogative clauses in this register were in the form of rhetorical questions, posed by the author to make a specific point, or to be directly answered by him or her, as examples (50) through (52) illustrate:

- (50) *Dè fon ghrèin a thachair don ghealltanas a fhuair sinn ann an 1948 nuair a chaidh a ghealltainn dhuinn nach biodh dìth no deireas oirnn bho bhreith gu bàs?*
 'What beneath the sun happened to the promise that we received in 1948 when it was promised to us that we would not have no lack or loss from birth to death?'
- (51) *Dè ged a chanadh Richards no duine eile gur e fìor dheagh dhuine a bh' ann an Sellar 's gun fheuchadh iad ri sin a dhearbhadh?*
 'What if Richards or anyone else said that Sellar was a truly good man and they tried to prove that?'
- (52) *Mar eisimpleir, cia mhead duine cloinne a chuala sib' fhèin a' bruidhinn an deagh Ghàidhlig bho chionn ghoirid 's nach do ghabh sibh iongnadh dheth?*
 'For example, how many children have you heard speaking in good Gaelic recently without you being amazed at it?'

In contrast, examples (53)–(55) from conversation, are far more mundane:

- (53) 1> *ciamar a chaidh dhuibh an-diubh?*
 2> *ò cha robh i ach gu math rapach gu dearbh fhèin*
 1> how did it go for you today?
 2> oh it [the weather] was just nasty indeed
- (54) 1> *'s cò an latha a chunnaic sibh a' falbh i?*
 2> *och dìreach an-dè fhèin feasgar an-dè*
 1> and which day did you see her leave?
 2> oh just yesterday itself yesterday afternoon
- (55) 2> *Dadaidh a bheil thu a' deanamh cofaidh?*
 1> *chan eil ach trobhadh fhèin a-nuas nì thu fhèin cupa deis*
 2> Daddy are you making coffee?
 1> no but come down here you'll make a cup ready

For imperative clauses, there were significant pairwise comparisons between traditional narrative and all of the other registers, save fiction, as indicated in the arrows below:

Table 45: Sorted data: imperative clauses

4.49	Traditional Narrative
2.39	Fiction
1.30	Conversation
.58	Radio Interview
.54	Popular Writing
.22	Formal Prose
.00	Sports Reportage
.00	News Scripts

This result is out of line with the expectation that conversation would have the highest occurrence level of imperatives. Had other situations been represented in the conversational sub-corpus, such as meal preparing or strictly dinner table conversations – where requests or demands are frequent – one could imagine much higher levels in the conversation texts. However, imperatives are not the only way of requesting action on the part of an addressee and, in many cases, would be felt to be overly direct or impolite. In conversation, it is often the case that interrogative clauses are used in their stead:

- 2> *Dad an gearr thu am pie agam?*
 1> *Carson nach gabh thu gearr thu am pie agad fhèin?*
 2> Dad will you cut my pie?
 1> Why don't you ... cut your own pie?

In traditional narrative, dialogue is often between figures of high and low status, such as a king and a son, with speech reflecting the vertical nature of the relationship. Moreover, like the real world perhaps, kings and their like are portrayed as doing more important talking than low status figures. There is a lot of direction giving in these tales, connected to search and rescue themes. Also, the severity of the situations – war, encounters with the supernatural, and so on – along with the previously mentioned characteristics, lends itself to language that is rough and ready, projecting a more exciting and dynamic picture of events and discourse. The following examples depict these aspects of traditional narrative:

- (56) *“Cò air a dhìolas mi bàs m'athar?” ars' esan. “Fhalbh,” arsa Fionn, “agus dìol air a' mhuir-làn ud shìos e,” ars' esan.*
 “On whom will I revenge the death of my father?” he said. “Away,” said Finn, “and revenge it on the high tide down there,” he said.
 (57) *“tha trungaichean òir,” ars' esan, “agus trungaichean airgid agam-sa ann an uamha a's a' ghleann ud shuas agus leig air mo chois mi agus bidh iad uileag leat-sa.”*
 “I have trunks of gold,” he said, “and trunks of silver in a cave in the glen up there and let me get up and they'll be all yours.”
 (58) *“Cumaibh oirbh a' dh'iomall a' cheò,” as esan, “s na biodh eagal oirbh.”*
 “Keep going to the edge of the mist,” he said, “and don't be afraid.”

6.1.6 Copular Verbs

The copular verb, in Gaelic, is found in a number of constructions, the most frequent being clefts. Jang (1998) reports that a higher prevalence of copular verbs in Taiwanese tends to occur in more logical and persuasive registers.

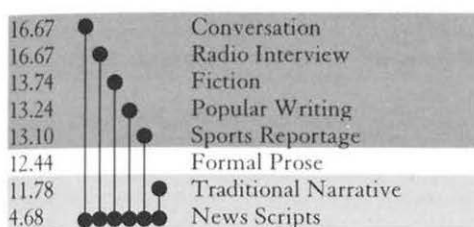
Data and commentary. A significant difference obtained for the registers in levels of copular verbs ($F = 6.427; p < .001$).

Table 46: Results: copular verbs

Register	Mean
Conversation	16.67
Radio Interview	16.67
Sports Reportage	13.10
Traditional Narrative	11.78
SPEECH	14.45
Fiction	13.74
Formal Prose	12.44
News Scripts	4.68
Popular Writing	13.24
WRITING	10.90
Total	12.67

As the arrows indicate, significant pairwise comparisons were found between news scripts (4.68) and all other registers besides formal prose:

Table 47: Sorted data: copular verbs



To a certain extent, it seems that Jang's results are applicable to this data. The news script register, being primarily concerned with factual information rather than persuasion, has the lowest occurrence. Conversation and radio interview, having the highest levels, could be construed as the most persuasive of the registers in the corpus, or at least the most concerned with expressing view-points and influencing others. However, a large question remains whether copulas in ScG are similar to those in Taiwanese. Clearly, further research is needed on the discourse functions of the copula in Gaelic.

6.1.7 Identificational Clefts

This construction is used for proper inclusion, to link a referent with a class to which it belongs (see Appendix 1). Example (59) illustrates this property:

- (59) *'s e croitear a th' ann an Ùisdean*
 'Hugh is a crofter'

Identificational clefts have not been investigated in previous register studies.

Data and commentary. A significant overall difference was found between the registers for identificational clefts ($F = 27.296$, $p < .001$) but no pairwise comparisons reached a significant level. The means are presented in Table 49:

Table 48: Results: identificational clefts

Register	Mean
Conversation	2.49
Radio Interview	2.53
Sports Reportage	3.91
Traditional Narrative	.88
SPEECH	2.45
Fiction	.94
Formal Prose	.29
News Scripts	.25
Popular Writing	1.21
WRITING	.68
Total	1.56

Without significant pairwise comparisons having obtained, it is impossible to state which registers are clearly different from one another on this feature. In a larger study, perhaps differences would be found, but as the current findings are inconclusive, this feature will not be discussed again in the context of this study.

6.2 Nominal Morphosyntax

6.2.1 Adjectives

A number of different adjective types were surveyed in this research: attributive adjectives; demonstrative adjectives; predicative adjectives; and superlative and comparative forms.

Some researchers report that **attributive adjectives**, or adjectives in general, are more frequently found in writing than speech (Drieman 1962; Chafe 1982; Biber 1995). Biber (1988) sees adjectives as one of the means of expanding and elaborating textual information. Miller and Weinert (1998) demonstrate how, in spontaneous spoken registers, speakers normally avoid complex noun phrases and stretch descriptions out over several clauses and phrases. An example of this tendency is the lower level of attributive adjectives in more oral registers. Biber (1988), Jang (1998), and Biber *et al.* (1999) state that attributive adjectives are characteristic of more informational registers, such as academic prose.

Demonstrative adjectives serve to specify a particular referent and distinguish it from other similar ones in a temporal-spatial field. They are important for maintaining textual cohesion (Halliday 1994). Ochs (1979) finds that, where planned texts would use definite articles, demonstratives are often preferred in unplanned ones (e.g. *this lady* [unplanned] *vs.* *the lady who* [planned]). Biber *et al.* (1999) find demonstrative adjectives to be most frequent in formal prose, but proportionately so in conversation where they are the most frequent type of definite determiner.

In Biber (1988), **predicative adjectives** were found to be weakly correlated with more abstract registers. Biber *et al.* (1999) find them slightly more common in fiction, where they most commonly indicate psychological states such as *afraid*, *happy*, or *tired*. Chafe (1982) and Biber (1988) state that attributive adjectives are integrative, whilst predicative adjectives are more fragmented.

Superlatives and comparatives are used to place the level of a certain quality possessed by a referent in relation to the levels possessed by other specified referents or a population. If their distribution is similar to attributive adjectives, we would expect them to be indicative of more integrative, more literary registers.

Data. The means for registers on the adjectival features are listed below:

Table 49: Results: adjectives

Register	Attributive	Demonst.	Predicative	Super/comp
Conversation	14.04	5.05	12.25	1.55
Radio Interview	15.49	4.50	13.36	3.08
Sports Reportage	15.15	1.15	4.98	1.54
Traditional Narrative	13.38	2.58	5.99	3.87
SPEECH	14.46	3.26	8.92	2.48
Fiction	28.03	2.31	12.60	1.63
Formal Prose	26.87	9.20	14.36	5.57
News Scripts	25.41	1.13	4.33	1.78
Popular Writing	25.40	4.93	14.43	1.99
WRITING	26.31	4.39	11.37	2.71
Total	20.39	3.82	10.15	2.59

There was a significant difference between the registers for **attributive adjectives** ($F = 8.75, p < .001$). A number of significant pairwise contrasts also obtained:

Table 50: Sorted data: attributive adjectives

28.03	●	Fiction
26.87	●	Formal Prose
25.41	●	News Scripts
25.40	●	Popular Writing
15.49	●	Radio Interview
15.15	●	Sports Reportage
14.04	●	Conversation
13.38	●	Traditional Narrative

As seen in the means at the left of Table 50, the written registers had, by far, a higher occurrence level of attributive adjectives than the spoken registers. There is a jump in ten occurrences per 1000 words between radio interview, at the top end of the spoken registers, and popular writing, at the bottom end of the written registers. Traditional narrative had the lowest level of occurrences in the corpus. Significant contrasts obtained between each of the spoken registers and each of the written registers, except for formal prose – which did not evince any significant comparisons.

There was also a significant difference in the registers for **demonstrative adjectives** ($F = 14.813, p < .001$).

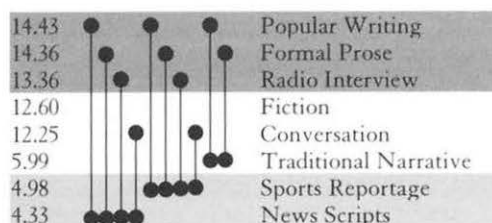
Table 51: Sorted data: demonstrative adjectives

9.20	●	Formal Prose
5.05	●	Conversation
4.93	●	Popular Writing
4.50	●	Radio Interview
2.58	●	Traditional Narrative
2.31	●	Fiction
1.15	●	Sports Reportage
1.13	●	News Scripts

Formal prose has the highest occurrence level for this feature, followed, after a significant gap, by conversation and popular writing. At the other end of the scale are sports reportage and news scripts, having roughly 1/9 the occurrences found in formal prose. Significant contrasts occurred between formal prose and most of the other registers, as well as between conversation and popular writing on the one hand, and sports reportage and news scripts on the other. Thus, there was a roughly three tiered distribution of contrasts, illustrated by the shading in Table 51 above.

With **predicative adjectives**, there was also a significant difference between the registers ($F = 8.492, p < .001$). Significant pairwise comparisons follow:

Table 52: Sorted data: predicative adjectives



Here, popular writing had the highest occurrence level of predicative adjective, followed by formal prose and radio interview. At the other end of the scale are sports reportage and news scripts, having the lowest returns for the feature. Significant contrasts, represented by the arrows, obtained between these two registers, along with traditional narrative *versus* popular writing, formal prose, and radio interview. There were also significant differences in conversation against sports reportage and news scripts. Examining the means, one notices a large drop between conversation and traditional narrative.

Finally, there was a significant difference between the registers for prevalence of **superlative and comparative adjectives** ($F = 3.673, p = .002$) but there were no significant pairwise comparisons. Without additional data, it is unclear to what extent the registers differ from one another on this variable. However, it is interesting to note that formal prose, which has the second highest return for predicative adjectives, also has the highest for superlatives and comparatives.

Commentary. In the attributive adjective results, the status of formal prose was slightly ambiguous as no significant contrasts obtained between it and the other registers; in past research it has consistently had a high occurrence level. However, this may be due to the high level of variance found; with a larger sample, it is reasonable to expect that statistical power would increase and a significant difference would obtain.

The main difference in the data for attributive adjectives is between the written registers, having a high occurrence of the feature, and the spoken registers, with lower occurrence levels. As will be seen in section 7.2 on complex noun phrases, there are substantial qualitative differences in the kinds and complexity of adjectives used in spontaneous speech and writing. While long stretches of attributive adjectives occur with some regularity in writing, speech generally prefers to spread descriptions out over several clauses. The reader is referred to the examples in section 7.2 for illustrations of this tendency.

For demonstrative adjectives, formal prose had the highest occurrence level, trailed, after a significant gap, by conversation and radio interview. Sports reportage and news scripts had the lowest returns for the feature. This is understandable given their strong forward motion; there is little returning in these registers to previous points of reference and a quick overturn of present ones.

In their role as a provider of textual cohesion, it is not surprising that demonstrative adjectives are most frequently found in formal prose. This register is bound up with the analysis and integration of separate strands of information and there is constant assimilation of preceding text. Although conversation has the second highest level of demonstrative adjectives after formal prose, it does not seem likely that it would use them in similar ways. Biber *et al.* (1999) found that the proximal forms of the demonstrative adjective (*this, these*) are more frequent in expository prose than the distal forms (*that, those*). On the other hand, they found that conversation had a greater proportion of distal forms. This is also the case in Scottish Gaelic, as can be seen in Table 53:

Table 53: Proportion of demonstrative adjectives in formal prose and conversation

Demonstrative Adj.	Formal Prose		Conversation	
	Count	Frequency	Count	Frequency
seo (proximal)	39	0.55	4	0.10
sin (near distal)	26	0.37	31	0.78
(si)ud (far distal)	6	0.08	5	0.13
Total	71	1.00	40	1.01

Formal prose has a marked preference for the use of *seo* 'this' while, in conversation, there is an even more marked preference for using *sin* 'that'. The use of *(si)ud* 'yon'³⁷ is of a similarly low occurrence in both registers, but slightly more frequent in conversation. This difference in distribution correlates with the differences in function evident between the two registers. Authors of formal prose tend to use demonstrative adjectives endophorically, referring backwards or forwards to referents within the text itself. The following examples illustrate this tendency:

- (60) *Tha sinn an-diugh ga fhaicinn nas iomchaidhe 's nas buannachdaile bàrdachd na Gàidhlig a rannsachadh 's a mheas a-rèir an dualchais dham buin i 's mu choinneamh an t-suidheachaidh a tha ga toirt gu bith. Agus anns an rannsachadh seo, an coimeas ris gach seòrsa eile bàrdachd a tha againn, rinneadh tomhas mòr de dhearmad air a' bhàrdachd spioradail.* (FORM)

'We today see it more appropriate and fruitful to research and judge Gaelic poetry according to the tradition to which it belongs and the situation that gives rise to it. In this research, compared to every other kind of poetry we have, a great omission was made of work on spiritual poetry.'

- (61) *Bha ealain a' bhàird a' toirt cothrom do'n luchd-èisdeachd a dhol an compàrt cultarach; agus bha an compàrt sin a' daingneachadh an grèim air an ionad chultarach.* (FORM)

'The bard's art gave a chance to the listeners to join in cultural participation; and that participation galvanised their hold on the cultural institution.'

- (62) *'S e an dòigh anns a bheil an t-atharrachadh seo a' tighinn gu bheil barrachd sliochd ga fhàgail leis an fheadhainn sin a tha comasach air iad fhéin a dheanamh suas ris a' choimhearsnachd, ...* (FORM)

'It is the way in which this change comes that a greater lineage is left by those ones who manage to succeed with the environment, ...'

In example (60), *seo* 'this' acts as an anaphor, modifying its head noun *rannsachadh* 'research' in such a way that it becomes co-referent with a previous usage of the word. In example (61), the same process occurs, but the anaphor in this case is *sin* 'those'. In contrast to these examples, example (62) uses the demonstrative pronoun *sin* as a cataphor, pointing ahead to the referent defined by the following restrictive relative clause. As mentioned previously these are instances of text-internal reference. In

³⁷ *(Si)ud* is commonly used in Gaelic while 'yon' remains common for speakers of Scottish English.

conversational texts, on the other hand, there are many more cases of demonstrative adjectives being used for the purposes of situational or text-external reference, and time reference:

- (63) 4> *rinn mi cupa eile an-shin*
 3> *'g iarraidh feadhainn dhiubh siud?*
 1> *feuch dhomh aonan*
 4> I made another cup over there
 3> wanting some of those ones?
 1> give me one
- (64) 4> *dè an gàrradh?*
 2> *shios air chùl an taighe an-sheo na h-àiteachan sin uileag*
 4> what wall?
 2> down behind the house here all those places
- (65) *och bha i cho garbh 's thàinig an droch mheall a bha siudach oirnn rinn sinn suas ar n-inntinn an uair sin gur e gun rachadh sinn gu clèibh a bh' againn air fasgadh*
 'oh it was so rough and that bad shower that was there came down on us we made up our minds [at] that time that it was that we would go to creels that we had in shelter'

In example (63), the referent of the NP *feadhainn dhiubh siud* is in the immediate physical context – a bunch of biscuits; *siud* is being used situationally, to point to this referent. In example (64), there is no text-internal point of reference for *sin* – the listeners are expected to know exactly where speaker 2 is indicating from their common knowledge of the area. Example (65) is a case of time reference, which, although not completely text-external – it is interpreted as indicating a time coreferential with the inception of the event in the second clause – is an unusual usage in the formal prose texts.

To summarise the findings on demonstrative adjectives, a difference was found between the two registers exhibiting the highest occurrence level of them: formal prose and conversation. In formal prose, most references were text-internal and utilised the proximal form *seo*. In conversation, there were more text-external or situational references than in formal prose and a greater use of the distal demonstrative adjectives *sin* and (*si*)*ud*. The reason offered for the low levels found in news scripts and sports reportage is their strong forward motion. In fact, when looking at the instances of demonstrative adjectives in sports reportage, one finds that over half of them (7 out of 12 instances) are at lapses in the game, the points at which commentary occurs.

The results for predicative adjectives are interesting but difficult to interpret. Popular writing had the highest level of the feature followed closely by formal prose and radio interview. Fiction and conversation were next and there was only a difference of roughly 2 occurrences per 1000 words between conversation and popular writing. Between conversation and traditional narrative – the next register in descending occurrence level – there was a drop of over 6 occurrences per 1000 words. This grouping of registers, with a mixture of writing and speech, is not immediately forthcoming as to potential functional underpinnings and is somewhat contrary to the findings of other investigators. However, the fact that news scripts and sports reportage have the lowest return for predicative adjectives gives support to Biber's (1988) finding that they are more correlated with abstract registers; these two registers are the least abstract in the corpus.

6.2.2 Appositives

Appositives are defined by Trask (1993) as "[a] noun phrase which immediately follows another noun phrase of identical reference, the whole sequence behaving like a single noun phrase with respect to the rest of the sentence... [It] serves only to provide additional information" (p. 19). Biber *et al.* (1999) found that appositive noun phrases were

common in news and academic prose registers. Of those found in news, over 90% modified a proper noun, the majority of which had human reference. In academic prose, on the other hand, while a large proportion of appositives involved a proper noun, very few had human reference.

Data. A significant difference was found in the registers on levels of appositives ($X^2 = 48.283, p < .001$). There was also a significant difference between the modes ($X^2 = 33.222, p < .001$). News scripts had the highest prevalence, followed by popular writing and formal prose, with all spoken registers save sports reportage showing very little. Looking at the minimum and maximum for news scripts, it is clear how deviant this register is compared to the others. The data are presented in Table 54 below:

Table 54: Results: appositives

Register	Mean	Median	Minimum	Maximum
Conversation	.00	.00	.00	.00
Radio Interview	.23	.00	.00	1.83
Sports Reportage	.87	.45	.00	2.98
Traditional Narrative	.20	.00	.00	2.01
SPEECH	.33	.00	.00	2.98
Fiction	.83	.43	.00	2.90
Formal Prose	2.64	1.97	.00	7.72
News Scripts	7.27	6.98	1.54	20.26
Popular Writing	3.41	2.88	.00	6.37
WRITING	3.70	2.80	.00	20.26
Total	2.02	.43	.00	20.26

Commentary. As in Biber *et al.* (1999) there was a high concentration of appositives in news scripts and formal prose. Popular writing also figured strongly. In light of the modal differences found, it seems that this is a feature of generally written language production, and probably a manifestation of the fewer production constraints and the higher informational emphasis placed in writing.

In news scripts, appositives are usually linked with human-referent proper nouns. Their presentation is usually name + apposition phrase, but they can obtain in the opposite order as well:

- (66) *Gheall Prìomh Mhinisteir na h-Alba, Domhnall Dewar, ...* (NEWS)
 'The First Minister of Scotland, Donald Dewar, promised ...'
- (67) *Thuir Eddie Stiven, eòlaiche air sàbhailteachd mara, ...* (NEWS)
 'Eddie Stiven, an expert on marine safety, said ...'

In addition to being used with personal names, appositives in news scripts are also associated with the names of groups:

- (68) *Tha a' bhuidheann glèidhteachais – Friends of the Earth – air iarraidh ...* (NEWS)
 'The conservation group – Friends of the Earth – has demanded ...'

Finally, they are used to introduce terms that would be unknown without glossing:

- (69) *dh'fhaodadh banacraich an aghaidh a' ghalair bhradain – ISA ...* (NEWS)
 'A vaccination against the salmon disease – ISA – may ...'

In the popular writing texts, all of the previous usages were found. Another usage found, which is sometimes also present in news scripts – but was not located in the news texts of this corpus – is in glossing an unfamiliar Gaelic word or neologism:

- (70) *'S ann don Linne Dubh – Blackpool ...* (POP)
'It is to the Linne Dhubh – Blackpool ...'

In the formal prose texts of the corpus, there were not large differences to be seen in appositive usage from the other registers, but there were a few instances of glossing entities:

- (71) *Tha càirdeas an t-seòrsa eile – a' bhàrdachd nuadhasach – ris an tradisean nas caime, ...* (FORM)
'The relationship of the other kind – the modern poetry – to the tradition is more indirect ...'

Perhaps if the formal prose included in the corpus had been of a more technical nature, there would have been a greater variety in the kinds of glossing seen. Unfortunately, there are very few representations of this kind of discourse to be had in Scottish Gaelic.

6.2.3 Complex Noun Phrases

Miller and Weinert (1998) is possibly the fullest treatment of the different degrees of noun phrase complexity in spontaneous speech. Their research has indicated that noun phrases with more than one adjective are unusual in spontaneous spoken English and German. Also, as mentioned previously, they found that speakers tend to stretch out descriptions over many clauses or phrases rather than lumping together sets of modifiers. They achieve a high level of description in their data, scoring complexity differently depending on discrete combinations of features, but it was felt that replicating this part of their work for the present research would have been overly time-consuming. Noun phrases were, instead, tagged as 'complex' when they had two or more modifiers of the following kinds: attributive adjectives, participles (rare in Gaelic data), prepositional phrases, other NPs as part of a genitival relation, and relative clauses. A complex NP could be as unadorned as a noun followed by two attributive adjectives or as involved as having various numbers of all of the above modifiers.

Data and commentary. A significant difference obtained between the registers for complex NPs ($F = 7.085, p < .001$). The means are listed in Table 55:

Table 55: Results: complex noun phrases

Register	Mean
Conversation	2.36
Radio Interview	5.30
Sports Reportage	7.00
Traditional Narrative	5.09
SPEECH	4.92
Fiction	9.49
Formal Prose	15.86
News Scripts	14.81
Popular Writing	11.94
WRITING	13.11
Total	9.01

There were also a number of significant pairwise comparisons:

Table 56: Sorted data: complex noun phrases

15.86	●	Formal Prose
14.81	●	News Scripts
11.94	●	Popular Writing
9.49	●	Fiction
7.00	●	Sports Reportage
5.30	●	Radio Interview
5.09	●	Traditional Narrative
2.36	●●●	Conversation

Formal prose has the highest mean for the feature, followed by news scripts and popular writing. On the other end of the spectrum is conversation, with roughly one seventh the occurrence level of the top two registers. As indicated by the arrows, there were significant contrasts between, mainly, formal prose and news scripts on the one hand, and conversation on the other. Additional contrasts obtained between traditional narrative and news scripts, and between conversation and popular writing.

Overall, the results indicate that complex NPs are much more common in the more informational, written registers. All of the spoken registers had fewer cases of complex NPs with the lowest found in conversation. This result agrees well with the findings of Miller and Weinert. Section 7.2 expands on these findings and presents a number of relevant examples to illustrate the differences between formal prose and the spontaneous spoken registers. The reader is referred to this section for additional information.

6.2.4 Nouns: Total Nouns, Nonconcordant Nouns, Genitives and Indefinite Marking

Four features were considered in this section: the total number of nouns; the ratio of nonconcordant nouns to concordant ones; the number of genitive nouns; and the frequency of indefinite marking.

6.2.4.1 Total Nouns

Biber (1988; 1995; *et al.* 1999) and Jang (1998) have linked high noun counts to written, informational registers such as academic prose and news. Conversation, on the other hand, has been found to have a low proportion of full nouns, but a high concentration of pronouns.

Data and commentary. There was a significant difference in the total number of nouns of the registers ($F = 39.639$, $p < .001$). The means are presented in Table 57:

Table 57: Results: total nouns

Register	Ns: total
Conversation	117.26
Radio Interview	138.39
Sports Reportage	203.40
Traditional Narrative	143.62
SPEECH	151.31
Fiction	195.34
Formal Prose	210.28
News Scripts	251.55
Popular Writing	212.43
WRITING	218.62
Total	184.97

The significant pairwise contrasts follow:

Table 58: Sorted data: total nouns

251.55		News Scripts
212.43		Popular Writing
210.28		Formal Prose
203.40		Sports Reportage
195.34		Fiction
143.62		Traditional Narrative
138.39		Radio Interview
117.26		Conversation

If noun counts are considered an index of informational richness, the news register is by far the most informational of the registers followed by, after a significant gap, popular writing and formal prose. On the other end of the spectrum, conversation has a low noun occurrence; less than half that of news scripts. As indicated in the arrows and shading, the main contrast distinguished in Table 58 above is between conversation, radio interview, and traditional narrative on the one hand, and news scripts on the other. There is another group of registers, which are in a medial position between news scripts and those just mentioned. These are more noun rich than most of the spoken registers (sports reportage is in the medial group), but less so than news scripts.

To illustrate the way in which conversation and news scripts differ in terms of how they encode meaning, the following excerpts have had all of their full nouns removed. As can be seen, this has a detrimental effect on the comprehensibility of both samples, but it is far more pronounced in the news script:

Example:

1> *ciamar a chaidh dhuibh an-diugh*

2> *o cha robh i ach gu math rapach gu dearbh fhèin o fhuair sinn rèidh air [] math dhiubh ach b' fheudar dhuinn am fàgail bhuainn mu dheireadh thall och bha i cho garbh 's thàinig an droch [] a bha siudach oirnn rinn sinn suas ar n-[] an uairsin gur e gun rachadh sinn gu [] a bh' againn air [] ach bha an [] cho ìosal cha d' fhuair sinn an togail ann 's dè bha a' dol agad fhèin?*

1> *o bha mi shuas ann am [] aig na [] 's thadhal mi air [] tha m' [] goirt 's tha mi a' call mo [] an seo-ach*

2> *o chan e droch [] tha sin idir gheibh sinn [] a-nochd co-dhiubh*

1> *cha bhi thu cho bragail ma dh' fheumas tu a bhith deanamh a h-uile [] a-staigh ma bhios mise bochd ge-tà*

Translation:

1> how did you get on today?

2> oh it was just nasty indeed oh we got settled on a good [] of them but we had to leave them from us finally och it was rough and that bad [] came on us and we made up our [] then that it was that we would go to [] that we had on [] but the [] was so low we didn't manage to lift them there and what was doing with you?

1> oh I was up in [] at the [] and I visited [] my [] is sore and I am losing my [] here

2> oh that's not a bad [] at all we'll get [] tonight anyway

1> you won't be so cheeky if you have to do every [] inside if I am ill however

Example:

Tha [] beag ceithir [] a dh'[] a' tighinn air adhart 'san [] am [] [] as dèidh [] a thoirt air le []. Chaidh [] [] a ghortachadh mun [] 's mun []. Chaidh an [] [] a chur gu [] le []. Tha [] air a' [] a' dol gu [] [].

Tadhlaidh [] [] [] air dà [] ann [] [] a tha 'g iarraidh [] agus [] a leasachadh. Cuiridh [] [] [] air dòigh ann an [] [] air an t-siathamh-[]-deug dhen ath [] gus an cluinn iad na tha dhith air [] an []. Ron a' [] chi iad [] [] far am bheil a' [] [] ag iarraidh an t-seann [] a leasachadh airson [] agus [] bheaga. Thèid iad cuideachd a [] far am bheil [] [] tur ùr air iarraidh.

Translation:

A [] four [] of [] is coming along in the [] in after an [] on him by a []. The [] was injured about the [] and the []. The [] was put to [] by a []. A [] on the [] is going to [] [].

[] will visit two [] in [] [] who are wanting to expand a [] and a []. The [] of [] is going to run the [] in [] [] on the seventeenth [] of the next [] so that they can hear what the [] of the [] are lacking. Before the [] they will see [] [] where the [] [] is wanting to develop the old [] for [] and small []. They will also go to [] where a completely new [] has been requested.

6.2.4.2 Grammatical Nonconcordance

Nonconcordance was defined as the presentation of a word form out of line with the conditions prescribed in the reference grammar (Appendix 1). Decisions about concordance or nonconcordance were based upon a set of predetermined paradigms as well as the lexical forms offered in those dictionaries consulted (see §4.3.1 and §4.3.2) for information on procedures). An example of such a form, utilising the previous example, would be *iuchair an taigh*, which presents the nominative form of 'house' where one would expect the genitive *taighe*. Nonconcordant noun forms have been regularly cited in Gaelic linguistic literature as indicative of low formality (see references in §2.4).

Data and commentary. There was a significant difference between the registers for the proportion of nonconcordant nouns to concordant nouns ($F = 12.70, p < .001$). The means are presented in Table 59:

Table 59: Results: grammatical nonconcordance

Register	Conc Ratio
Conversation	.03
Radio Interview	.06
Sports Reportage	.10
Traditional Narrative	.02
SPEECH	.05
Fiction	.02
Formal Prose	.01
News Scripts	.06
Popular Writing	.03
WRITING	.03
Total	.04

The significant pairwise comparisons follow:

Table 60: Sorted data: grammatical nonconcordance

.10	● ● ● ●	Sports Reportage
.06	● ● ● ●	Radio Interview
.06	● ● ● ●	News Scripts
.03	● ● ● ●	Conversation
.03	● ● ● ●	Popular Writing
.02	● ● ● ●	Traditional Narrative
.02	● ● ● ●	Fiction
.01	● ● ● ●	Formal Prose

Sports reportage had the highest proportion of noun nonconcordance of all of the registers, followed by radio interview (nonsignificant pairwise comparisons) and then news scripts. On the other end of the scale are formal prose, fiction, and traditional narrative, which had the least nonconcordance of any spoken register. It is surprising that news scripts, which one could reasonably take to be amongst the more conservative registers in modern Scottish Gaelic, would have such a high proportion of nonconcordant forms. On the other hand, it is not surprising that formal prose and fiction evince conservative noun formation, as they are the most edited registers in the corpus. Additional information on concordance can be found in §7.1 below, which closely examines the registers for their treatment of noun and adjective case marking in the nominative, dative, and genitive, and provides numerous examples.

6.2.4.3 Genitives

Chafe (1982) indicates that genitives occur more often in writing and are characteristic of integrative styles. In ScG, genitives are indicated through word order and, in more conservative usage, case marking. For instance, in the Gaelic translation of 'the key of the house', 'the key' is in nominative case while 'the house' follows it and receives genitive case marking:

- (72) iuchair an taighe
 key-N ART house-G (cf. *taigh*, nominative)
 'the key of the house'

For more information on case marking in ScG, refer to §2.2.1.4 in Appendix 1.

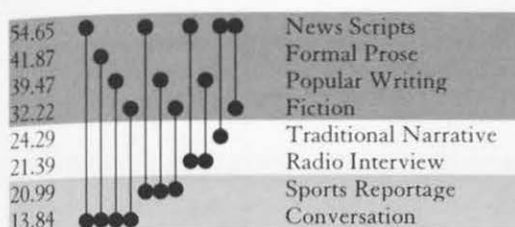
Data and commentary. A significant difference obtained for the prevalence of genitive nouns (i.e. in genitival positions – not necessarily morphologically genitive) between the registers ($F = 14.714$, $p < .001$). The means are as follows:

Table 61: Results: genitive nouns

Register	Genitive Ns
Conversation	13.84
Radio Interview	21.39
Sports Reportage	20.99
Traditional Narrative	24.29
SPEECH	20.06
Fiction	32.22
Formal Prose	41.87
News Scripts	54.65
Popular Writing	39.47
WRITING	42.51
Total	31.28

Post-hoc procedures revealed the following significant contrasts:

Table 62: Sorted data: genitive nouns



Clearly, news scripts have the most genitive nouns and all written registers have more than spoken ones. Conversation has the lowest incidence of the feature, with traditional narrative having the highest for a spoken register. There are four times the amount of genitive nouns in news scripts than there are in conversation. The arrows indicate significant contrasts between, mainly, all of the written registers, and conversation and sports reportage. The occurrence of genitive nouns seems to be strongly related to mode. For examples of genitive constructions and further discussion, please refer to §7.1.3 and §7.2.3.

6.2.4.4 Indefinite Marking

Indefinite marking of nouns in ScG is managed by the phrases *air choreigin* 'or another' (shortened to *-eigin* as in *duineigin* 'some man or another') and *sam bith* 'at all' (lit. 'in existence') as in *duine sam bith* 'any person at all/anyone'. Although their connotations are separate, their both serve to reduce the specificity of a referent and may be indicative of a more fragmented style. Tannen (1980) indicates that speech characteristically uses phrases such as 'some young guy...' while written registers prefer indefinite articles, e.g. 'a young person...'.

Data and commentary. Table 63 summarises the means for indefinite marking:

Table 63: Results: indefinite marking

Register	Indefinite Marking
Conversation	4.99
Radio Interview	4.01
Sports Reportage	2.33
Traditional Narrative	2.64
SPEECH	3.47
Fiction	3.50
Formal Prose	1.27
News Scripts	1.50
Popular Writing	3.11
WRITING	2.33
Total	2.90

There was a significant difference between the registers for this feature ($F = 3.781$, $p = .002$). Significant pairwise comparisons follow:

Table 64: Sorted data: indefinite marking

4.99	●●	Conversation
4.01	●●	Radio Interview
3.50	●●	Fiction
3.11	●●	Popular Writing
2.64	●●	Traditional Narrative
2.33	●●	Sports Reportage
1.50	●●	News Scripts
1.27	●●	Formal Prose

Conversation has the highest occurrence of indefinite pronoun marking and is followed by radio interview and fiction. At the lower end are news scripts and formal prose. As indicated by the arrows, there are significant differences between conversation and these two registers. There seems to be a definite trend for more information-oriented registers to have fewer instances of this feature. On the other hand, involved registers – such as conversation and radio interview – tend to have more. The use of indefinite marking serves to decrease the exactitude of a reference. Where specificity is required, as in information-based registers, this kind of vagueness is normally considered infelicitous or sloppy. Some examples from the registers follow below as illustrations:

- (73) *a bheil an cnatan air duine sam bith eile thall an-sin a Dhòmhnaill?* (CONV)
 'does anyone else have the cold over there Donald?'
- (74) *chuala mi Eachann a' ràdh uaireigin gun robh uaireigin a bhiodh athair o chionn fada a' reic èisg mun cuairt fhios agad air a' chladach an-shin* (CONV)
 'I heard Hector saying one time or another that there was one time or another that his father was long ago selling fish around you know on the shore there'
- (75) *... chan e idir seann bhrògan sam bith a tha 'n seo. Is e a th' ann ach brògan Choinnich, agus brògan Choinnich a-mhàin.* (FICT)
 '... it isn't at all just any old shoes here. What it is is Kenneth's shoes, and Kenneth's shoes only.'
- (76) *tha cuideigin òg no dhà eile a thig a-staigh ann 's a chumas a' dol a' ghnothach* [sic] (INTER)
 'there is a young person or another or two others who come in and keep things going'

6.3 Possession, prepositions, and pronouns

6.3.1 Possession

Two types of possessive constructions were surveyed: those using possessive pronouns (simple possession) and those enlisting the preposition *aig* 'at' (periphrastic possession). In ScG, there is an old, iconic distinction between inalienable nouns, which are possessed using possessive pronouns, and alienable ones, which are possessed using, generally, *aig*. Biber *et al.* (1999) find that possessive pronouns are more frequent in conversation and fiction than in news and academic prose.

Table 65: Results: possession

Register	<i>Aig possession</i>	<i>Simple possession</i>
Conversation	17.56	8.81
Radio Interview	13.18	9.92
Sports Reportage	11.24	23.25
Traditional Narrative	9.93	18.97
SPEECH	12.97	15.52
Fiction	9.98	30.61
Formal Prose	8.09	14.68
News Scripts	3.70	9.47
Popular Writing	9.02	16.62
WRITING	7.60	17.22
Total	10.28	16.37

Data and commentary. A significant difference obtained between the registers for *aig* possession ($F = 6.741, p < .001$). The pairwise comparisons follow :

Table 66: Sorted data: *aig* possession

17.56	●	Conversation
13.18	●	Radio Interview
11.24		Sports Reportage
9.98		Fiction
9.93		Traditional Narrative
9.02		Popular Writing
8.09	●	Formal Prose
3.70	●	News Scripts

Table 66 indicates that the conversational registers have the highest occurrence of the feature while formal prose and news scripts have the lowest. Significant contrasts obtained between conversation and radio interview, on the one hand, and these two written registers. There seems to be a distinction between more involved registers, which have more *aig* possession, and more information-based ones, which evince less. At first glance, this could be related to the greater number of pronouns, in general, in the conversational registers.

There was also a significant overall difference between the registers for *simple* possession ($F = 15.596, p < .001$). Many pairwise contrasts were also significant:

Table 67: Sorted data: *simple* possession

30.61	●	Fiction
23.25	●	Sports Reportage
18.97	●	Traditional Narrative
16.61	●	Popular Writing
14.68	●	Formal Prose
9.92	●	Radio Interview
9.47	●	News Scripts
8.81	●	Conversation

Table 67 describes a much different distribution pattern than Table 68. Fiction has the greatest frequency of simple possession followed by sports reportage. On the other end of the scale are conversation, news scripts, and radio interview. Significant contrasts

obtained between, mainly, these registers mentioned, as indicated by the arrows and the shading above.

These results are particularly interesting in that, although to English-speaking minds possession is a relatively clear-cut issue, the two forms in ScG used to communicate it are widely divergent in their distribution. Conversation ranks the highest for use of periphrastic possession but the lowest for simple possession. Radio interview also seems to switch its place in the two tables. Fiction, which is ranked in the middle in terms of periphrastic possession, has the highest occurrence of simple possession. There seems to be a complex interplay of sociolinguistic and grammatical influences at work here and more research into the discourse attributes of the two constructions is needed to determine the reasons for the variation. One possible train of investigation would be to consider the lexical differences involved in governing the choice of construction. This would determine whether current usage maintains the old alienable *vs.* inalienable contrast. Another would be to examine the possibility that the *aig* construction has become the standard way of indicating possession in modern Gaelic and that simple possession is mainly preserved in more conservative registers such as fiction. Although sports reportage has the second highest return for simple possession above, a closer consideration of the data indicates that most of the 'hits' were from a very small number of lexical items, almost all being body parts. Removing this register from the list above would leave only the written registers in the top half of Table 67. This gives credence to the possibility that simple possession is a grammatically conservative construction, and is being displaced by a periphrastic form. There seems to be a general movement towards periphrasis in the modern language, as seen in the marking of passive voice (§6.1.3) and certain other features (see §3.5 in Appendix 1).

Some examples follow from the various registers:

Aig possession

- (77) *uill dh'aithnich mi i chan eil an t-ainm aice agam idir* (CONV)
'well I recognised her I don't have **her** name at all' [Lit. 'the name **at her** is not at **me** at all']
- (78) *uill tha mi uabhasach toilichte a chluinntinn gu bheil na sgiobaidhean agaibh cho làidir* (INTER)
'well I am very pleased to hear that **your** teams are so strong'
- (79) *Tha an gine son seòrsa fala A2 ri fhaotainn mar as trice anns an Roinn-Eòrpa a-mhàin, agus chan eil i aig an t-sluagh Mhongolach.* (FORM)
'the gene for the blood-type A2 is found most often in Europe only, and the Mongolian people don't have it' [Lit. 'it is not **at the** Mongolian people']
- (80) *tha am ball aig Yugoslavia air taobh thall na pàirce* (SPORTS)
'the ball is **at** Yugoslavia on the other side of the pitch'

Simple Possession

- (81) *'S mu dheireadh thall, fhuair i a h-iartras, ...* (FICT)
'And finally, she got **her** wish, ...'
- (82) *bha uachdarain agus an luchd gnothaich a' gèilleadh dhan bheachd...* (FORM)
'landowners and **their** business people were yielding to the opinion ...'
- (83) *fhuair Miller a chas thuice an toiseach* (SPORT)
'Miller got **his** foot to her [the ball] first'
- (84) *ach bhiodh iad air an toir fairis feadhainn dhiubh* (CONV)
'but they would be exhausted some of them [Lit. '...would be after **their** taking past']

As can be seen from the examples, one difference between periphrastic and simple possession is that the periphrastic form can be used both predicatively, as in examples (79) and (80), and attributively, as in example (78). Example (77) shows both uses. Simple

possession, in contrast, can only be used attributively. Both forms of possession are used in a number of expressions that would not translate as possession in English. An example of this is *thèid agam air a dhèanamh* 'I can manage it', literally 'will go at me on its doing'. Both forms are present in this idiom – *aig* in *agam* and the 3rd person masculine pronoun *a*. For more information on possession in Scottish Gaelic, refer to §2.2.1.5 in Appendix 1.

6.3.2 Prepositions

Simple, complex, and compound prepositions were combined to compose this feature. Also included were prepositions integrated within noun phrase constituents. Jang (1998) finds that prepositions occur most frequently in more informational registers. Chafe (1982) describes prepositional phrases as a way of compressing information into an idea unit while Chafe and Danielewicz (1986) mention that prepositions are more likely to occur linked with verbs in conversation. Biber (1988) finds that prepositions correlate with noun counts, word lengths, attributive adjectives, and lexical variety and tend to be most common in typically literate, informational texts. Miller and Weinert (1998) discuss prepositional phrases in terms of noun phrase complexity, and find them to be rare in spontaneous spoken texts.

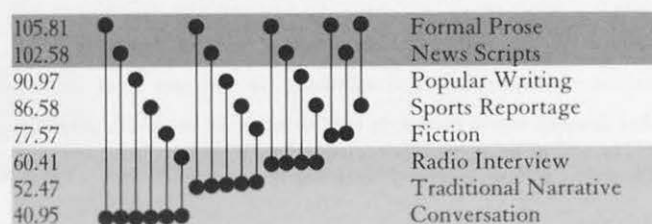
Data and discussion. A statistically significant difference obtained across the Scottish Gaelic registers for prepositions ($F = 37.438, p < .001$).

Table 68: Results: prepositions

Register	Mean
Conversation	40.95
Radio Interview	60.41
Sports Reportage	86.58
Traditional Narrative	52.47
SPEECH	60.09
Fiction	77.57
Formal Prose	105.81
News Scripts	102.58
Popular Writing	90.97
WRITING	94.72
Total	77.40

As can be seen in Table 69, there were many statistically significant contrasts in the pairwise comparisons:

Table 69: Sorted data: prepositions



Formal prose and news scripts had the highest occurrence levels for the feature with conversation, traditional narrative, and radio interview having the least. There were significant contrasts mainly between these groups, which seemed to indicate a strong modal factor, specifically a difference between information-rich, written registers and more typically oral ones. This result is in line with previous research.

Sports reportage, which has the highest occurrence of prepositions in the spoken texts, may be a manifestation of the association between prepositions and verbs, as reported by Chafe and Danielewicz. If we consider the occurrence of verbal nouns per 1000 words in the registers, we find the following:

Table 70: Results: verbal nouns per 1000 words

<i>Register</i>	<i>Mean</i>
Conversation	30.78
Radio Interview	36.24
Sports Reportage	49.36
Traditional Narrative	24.11
SPEECH	35.12
Fiction	26.83
Formal Prose	20.96
News Scripts	31.71
Popular Writing	27.63
WRITING	26.78
Total	30.95

Sports reportage clearly has the highest occurrence of verbal nouns in the corpus. A quick examination of the concordances from the verbal noun tags shows that around 1/3 of these returns are associated with a preposition. It seems reasonable to suggest that registers involving description of active movement have a higher percentage of verb + preposition constructions than other registers.

6.3.3 Pronouns

Two types of pronouns were investigated: personal pronouns and demonstrative pronouns. Personal pronouns provide anaphoric reference to previously mentioned active or retrievable NPs, or, in the case of 1st and 2nd person pronouns, they refer to context-retrievable addressers and addressees. Biber *et al.* (1999) found that pronouns in general are most frequently encountered in conversation. The tag scheme of the current research did not differentiate between the different person-types of the pronouns. On demonstrative pronouns, Biber (1988) writes:

[They] can refer to an entity outside the text, an exophoric referent, or to a previous referent in the text itself. In the latter case, [they] can refer to a specific nominal entity or to an inexplicit, often abstract, concept (e.g., *this shows ...*) (p. 226).

Biber's results indicated that demonstrative pronouns tend to occur in involved, generally oral texts (his Dimension 1) as well as being weakly correlated with on-line, informational registers such as prepared speeches and interviews (Dimension 6).

Data and commentary. Table 71 lists the means for the various registers on demonstrative pronouns and total pronouns:

Table 71: Results: pronouns

Register	Demonstrative PN	PNs: total
Conversation	9.52	67.69
Radio Interview	16.35	51.39
Sports Reportage	6.84	53.77
Traditional Narrative	10.56	108.52
SPEECH	10.53	71.34
Fiction	5.71	61.83
Formal Prose	5.70	24.50
News Scripts	2.50	18.27
Popular Writing	5.14	33.93
WRITING	4.70	33.45
Total	7.61	52.39

There was a significant difference between the registers for both features:

Demonstrative PNs	$F = 11.016$	$p < .001$
PNs total	$F = 55.871$	$p < .001$

Demonstrative pronouns showed the following significant pairwise comparisons:

Table 72: Sorted data: demonstrative pronouns

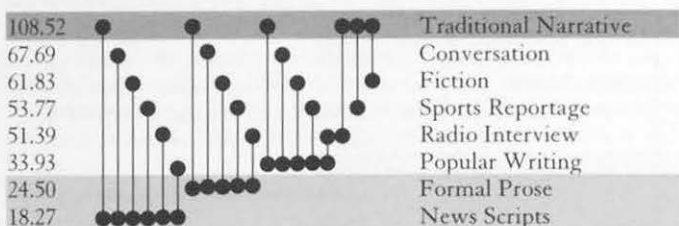
16.35	●	●	●	●	Radio Interview
10.56	●	●	●	●	Traditional Narrative
9.52	●	●	●	●	Conversation
6.84	●	●	●	●	Sports Reportage
5.71	●	●	●	●	Fiction
5.70	●	●	●	●	Formal Prose
5.14	●	●	●	●	Popular Writing
2.50	●	●	●	●	News Scripts

Radio interview clearly has the greatest prevalence of demonstrative pronouns, followed by traditional narrative and conversation. News, at the other end of the scale, evinced a very low level of the feature. The strongest contrasts, as indicated by the arrows, were between radio interview and all of the written registers. There were also statistically significant contrasts between news scripts, on the one hand, and traditional narrative and conversation on the other.

In general, the oral mode seems to correlate positively with a greater frequency of demonstratives. This distribution agrees to a large extent with Biber's findings. Perhaps the reason that news scripts have the lowest occurrence, is due, once again, to the fact that they seldom sum-up or refer to previous referents but, instead, manage a steady change of topic.

The total number of pronouns found contrasted significantly in the registers as follows:

Table 73: Sorted data: total pronouns



This distribution is similar to the previous one, with traditional narrative and conversation exhibiting some of the highest occurrences of pronouns while formal prose and news scripts have some of the lowest. The means indicate a wide range of occurrence level, with news scripts having roughly 1/6th the frequency of traditional narrative for pronouns. There were significant differences between many of the registers, with the most prominent differences being between news scripts and formal prose, on the one hand, and traditional narrative on the other.

Overall, those registers with a greater need to track participants have a greater occurrence of pronouns in their texts. It is not surprising that traditional narrative, being monologue and an extended account of usually multiple referents, would have, by far, the highest occurrence of pronouns. In addition, there seems to be an general tendency for more narrative-based texts to have a greater level of pronouns: fiction is written narrative and conversation, as has already been discussed, often features embedded narrative discourse.

By examining a traditional narrative excerpt, it becomes clear how important pronouns are for managing reference in this register:

Example Traditional narrative excerpt:

Well 's ann mar seo-ach u bhà, dh' fhalbh an Eachlair Ùrlair. Agus larna-mhàireach a seo bha an rìgh bhiodh e a' falbh a's a' bheinn-sheilg daonnan agus nuair a dh' fhalbh e dhan bheinn-sheilg co-dhiubh thuir i ri am fear a b' òige dhe na gillean thuir i ris, "a' ghraidheim," ar's esan ar's ise, "nach rathaid tusa suas," ar's ise, "agus, tha, chon an Eachlair Ùrlair," ar's ise, "a dh' iarraidh na Cìre Buidhe," ar's ise. "Agus cha bhì agad ri ràdh rìthe," ar's ise, "ach gun tàinig tu nuas, gun do chuir do mhuipe a nuas," ar's ise, "a dh' iarraidh na Cìr Buidhe thu." Dh' fhalbh e, an gille bochd ghabh e suas rànaig e taigh an Eachlair Ùrlair.

Translation:

Well, this is the way it was, the Eachlair Ùrlair left. And this next day the king, he would go off to the hunting-mountain always and when he left to the hunting-mountain anyway she said to the youngest one of the boys she said to him, "ah love," he said she said, "wouldn't you go up," she said, "and – up – to the Eachlair Ùrlair," she said, "to get the Yellow Comb," she said. "And there will be nothing for you to say to her," she said, "but that you came up, that your stepmother sent you up," she said, "to get the Yellow Comb." He left, the poor lad he went up and he reached the house of the Eachlair Ùrlair.

In this short piece of one hundred and three words, there are 21 pronouns, referring to the king, the queen, the youngest son, and the Eachlair Ùrlair; they comprise nearly 1/6 of the total lexical items (N=122). Not only do they indicate which referent is activated, but they are also important for reinforcing cues about which perspective is being taken, between narrative and direct quotes.³⁸ For instance, when the queen is speaking to the son, he is referred to using a 2nd person form, but when the narration resumes, he is referred to with a 3rd person masculine form.

¹⁰ Spoken language does not have quotation marks and hearers must rely on other types of cues to determine the type of discourse being used. These texts came from transcriptions made by others, thus punctuation was left in.

Pronominal forms are also important in traditional narrative, as in other registers featuring interaction, for the coding of status. As mentioned in §2.2.2 in Appendix I, Gaelic has a limited honourific system, roughly parallel to that of French and German. The following examples demonstrate honorific shift in a narrative tale ('Conal Gulban'). In example (85), King Conal is speaking to a supernatural being, who addresses him using the familiar form *thu*. In (86), he is being addressed by his subjects, who use the polite form *sibh*:

- (85) *'s bha am bodach a' tighinn as a dheoghaidh 's a bhoineid 'na dhòrn. "O till, till," ars' esan, "a Rìgh Eirinn. Chan ann mar sin a tha thu 'dol a dh'fhalbh," ars' esan, "agus pàisd' aig a' nighinn ruaidh air a bhreith dhut" ars' esan.*
'and the old man was coming after him clutching his bonnet. "Oh come back, come back," he said, "King of Ireland. It's not like that that **you** [familiar] are going to leave," said he, "with a baby having been born to **you** [familiar] by the red maiden," he said.
- (86) *Agus a chuile duine a choinneachadh a' Rìgh rachadh e air a ghluin dha 's dh' fhoighneachadh e dheth, "O cà robh sibh a Rìgh Eireann, o chionn trì ràithean?"*
'And every person who met the King, he would go down on his knee to him and he would ask him, "Oh where were **you** [polite] King of Ireland, for three seasons?"

6.4 Adverbs

Three different features were considered in this section: place adverbs; time adverbs; adverbs in general (*not* time or place). Jang (1998) found that counts of general adverbs were higher in narrative, while Biber (1988) discovered that they tended to be, along with place and time adverbs, most prevalent in registers such as broadcast and conversation, which, he says, are on-line and feature prominent use of exophoric reference. Narrative was found by Biber (1988) to have the highest instance of place adverbs, and by Jang (1998), to have the highest occurrence of time adverbs

Data and commentary. The mean differences are reported in Table 74:

Table 74: Results: adverbs

Register	Place	Time	Other Adv's
Conversation	18.60	13.97	15.64
Radio Interview	9.11	13.15	11.77
Sports Reportage	50.80	13.55	13.12
Traditional Narrative	22.14	11.65	9.92
SPEECH	26.00	13.08	12.66
Fiction	9.51	11.68	11.20
Formal Prose	3.99	6.31	11.30
News Scripts	1.97	13.71	12.01
Popular Writing	4.05	11.10	11.12
WRITING	4.64	10.77	11.41
Total	15.32	11.92	12.04

There was a significant difference between the registers for **place adverbs** ($F = 79.329, p < .001$). Pairwise comparisons obtained as follows:

Table 75: Sorted data: place adverbs

50.80	●	Sports Reportage
22.14	●	Traditional Narrative
18.60	●	Conversation
9.51	●	Fiction
9.11	●	Radio Interview
4.05	●	Popular Writing
3.99	●	Formal Prose
1.97	●	News Scripts

Sports reportage had the most place adverbs by far, with the next highest register – traditional narrative – having less than one half its occurrence level. Conversation also had a relatively high return for the feature. On the other hand, news scripts had a very low return – 4% that of sports reportage. Formal prose and popular writing also had low levels of place adverbs. The most prominent contrast was between sports reportage and all other registers. There were additional significant contrasts between, generally, the bottom three registers (news scripts, formal prose, and popular writing) and the top three (sports reportage, traditional narrative, and conversation).

Here, it is clear that registers placing a high demand on extra-textual reference for establishing context have higher counts of place adverbs. At the very top is sport reportage, in which addressers must convey to the audience a wide variety of spatial information as players interact with each other and the ball. Biber (1988) and Jang (1998) found that narrative-rich registers had the highest occurrences of place adverbs in their corpora. This would have been the case here as well, had sports reportage not been included. At the lower end of the spectrum are the more abstract registers, which do not have as great a need for this kind of reference. Place adverbs in such texts, especially those intended for readership (e.g. formal prose) rather than a listening audience (e.g. news scripts), sometimes show instances of intra-textual reference, coded in statements such as 'as we mentioned above'. Some examples follow illustrating the difference between the extra-textual use of place adverbs in sports reportage and traditional narrative, and the intra-textual use of them in formal prose:

Extra-textual

- (87) *Thàinig Gough a-steach air an taobh thall ann an-siud gun duine toirt sùil sam bith air* (SPORT)
 'Gough came in on the other side over there without anybody giving a glance at all at him'
- (88) *Ach 's e am plan' a rinn Fionn gu fàgadh e 'staigh fear air a robh Garaidh ann am falach a' watch-adh nam boireannach* (TRAD)
 'But it is the plan that Finn came up with that he would leave a man named Gary hidden inside watching the women'

Intra-textual

- (89) *Anns an àm ud, tha na Goill – na h-eachdraichean shuas mar eisimpleir ...* (FORM)
 'In that time, the Lowlanders – the historians above for example...'
- (90) *Anns an ruith thairis a th' agam an seo air puing no dhà, ...* (FORM)
 'In the quick run through I've done here on a point or two, ...'

In examples (87) and (88), the external reference is clear, relating to a place in the physical world depicted by the language in the text. In examples (89) and (90) however, the reference is to a place in the text. In (89), it concerns a prior point in the text while, in (90), it concerns the text as a whole.

A significant difference between the registers also obtained for **time adverbs** ($F = 2.795, p = .013$). The significant pairwise contrasts follow:

Table 76: Sorted data: time adverbs

13.97	Conversation
13.71	News Scripts
13.55	Sports Reportage
13.15	Radio Interview
11.68	Fiction
11.65	Traditional Narrative
11.10	Popular Writing
6.31	Formal Prose

Although conversation returned the highest occurrence for time adverbs, the news scripts register was the highest statistically significant one for this feature, followed closely by sports reportage. Formal prose deviated the most from the other registers, with an occurrence level that was less than half that of news scripts. Significant contrasts occurred between formal prose and three other registers: news scripts; sports reportage; and fiction.

The clearest contrast for time adverbs was between formal prose, and seemingly, all of the other registers, which had means that were very close to one another. Considering the previous result for place adverbs, formal prose seems to be less grounded to the physical world than the other registers; it is more abstract. In contrast, although news scripts had the lowest mean for place adverbs, it had the second highest mean for time adverbs. Similar to formal prose, the concerns of news scripts are generally uninvolved with physical space, but they are, in contrast, with temporal flow. In news scripts, there is a prominence of time adverbs such as *an-diugh* 'today', *a-màireach* 'tomorrow', *an uiridh* 'last year', *a-nochd* 'tonight', *an ath-sheachdainn* 'next week', *a-raoir* 'last night' and so on, indicating an anchoring of events to the future and past. Sports reportage also had a high return for time adverbs, but their reference differs somewhat. In this register, there is a greater instance of ones such as *a-nis* 'now (contrastive)', *a-rithist* 'again', *fhathast* 'yet/still', *mar-thà* 'already', *'sa bhad* 'instantly', and *mu dheireadh* 'finally', giving the impression that temporal reference is much more focused on the current state of affairs. This is not surprising given the different communicative functions of the two registers: news scripts concern a wide variety of past and present events, and possible future events; sports reportage is mainly concerned with the here and now, making only occasional references to past or future games.

There was no significant difference between the registers for **other adverbs** ($F = 1.136, p = .351$ n.s.). Judging by the distribution of the means (see Table 74 above), it is possible that additional data would confirm the tendency located by Biber (1988), for general adverbs to be prevalent in conversation and broadcast. It also seems to agree with Jang (1998), as narrative clearly has the lowest level of general adverbs. Despite these potential indications, it is not possible to be certain about the data, precluding further comments on this result.

6.5 Lexical Classes

This section involves a number of miscellaneous lexical categories between different adverb types, numerals, and verb types.

6.5.1 Adverbs of Stance and Degree: Amplifiers, Downtoners, Emphatics, and Hedges

Amplifiers are used to increase a speaker's conviction of a proposition. In ScG, words like *gu léir* 'completely', *direach* 'exactly', *glè* 'very', and *ro* 'extremely' serve this purpose. These differ from emphatics (see below) in that amplifiers give an index of the level of

certainty expressed. Amplifiers can also convey, as Holmes (1984, cited in Biber 1988) points out, agreement with a listener. **Downtoners** serve the opposite purposes of amplifiers, reducing conviction and conveying disagreement with a listener. One example in Gaelic is *gu ìre* 'to an extent'.

Emphatics, such as *gu dearbh* 'indeed', *gu deimhinne* 'with certainty', and *gun teagamh* 'without a doubt', express conviction towards a proposition without giving an indication of the degree of conviction. **Hedges** differ from downtoners in the same way that emphatics differ from amplifiers: they express a lowering of conviction towards a proposition without an indication of degree. Examples in Gaelic are *cha mhòr* 'almost', '*s dòcha* 'perhaps' and *math dh'fhaodte* 'perhaps'.

Several researchers have correlated the presence of these features with involved styles or the expression of personal stance (*amplifiers* Biber 1988; *downtoners* Biber 1995; *emphatics* Chafe 1982; Biber 1995; Jang 1998; *hedges* Biber 1988, 1995). Emphatics and hedges have also been linked with registers that are more typically oral than literate (*emphatics* Biber 1995; *hedges* Devito 1966, 1967; Chafe and Danilewicz 1987; Biber 1995).

Data. The means for these features are summed up in Table 77:

Table 77: Results: adverbs of stance and degree

Register	Amplifiers	Downtoners	Emphatics	Hedges
Conversation	5.12	.34	.96	1.26
Radio Interview	6.80	.24	3.34	3.01
Sports Reportage	7.74	.63	.95	1.98
Traditional Narrative	3.94	.28	.85	.88
SPEECH	5.85	.38	1.43	1.72
Fiction	5.36	.44	1.16	1.91
Formal Prose	5.18	.74	1.46	1.76
News Scripts	1.82	.05	.00	.51
Popular Writing	5.21	.31	.80	1.29
WRITING	4.34	.37	.82	1.33
Total	5.10	.37	1.13	1.52

There was a significant difference between the registers for **amplifiers** ($F = 4.065$, $p = .001$). The significant pairwise comparisons were as follows:

Table 78: Sorted data: amplifiers

7.74	●	Sports Reportage
6.80	●	Radio Interview
5.36		Fiction
5.21	●	Popular Writing
5.18		Formal Prose
5.12	●	Conversation
3.94		Traditional Narrative
1.82	●●●●●	News Scripts

Here, we see that sports reportage and radio interview have the greatest proportion of amplifiers while news scripts have the least. Significant pairwise comparisons, indicated by the arrows, obtained between these two groups as well as between news scripts, on the one hand, and popular writing and conversation on the other.

A significant difference also obtained for **hedges** ($F 3.229$, $p = .005$). The sorted data with significant pairwise comparisons follow:

Table 79: Sorted data: hedges

3.01	●●	Radio Interview
1.98	●	Sports Reportage
1.91	●	Fiction
1.76	●	Formal Prose
1.29	●	Popular Writing
1.26	●	Conversation
.88	●	Traditional Narrative
.51	●	News Scripts

Radio interview had the greatest occurrence of hedges while news scripts and traditional narrative had few compared to the rest of the registers. There was a significant difference between radio interview and these two registers. In terms of range, the mean occurrence for news scripts was about one sixth that of radio interview.

The occurrence levels of **emphatics** did differ significantly between the registers ($F = 3.953$, $p = .001$), but there were no significant pairwise comparisons. However, when considering the means (see Table 77), radio interview has a considerably higher amount than most other registers, while news scripts evinced none. Although post-hoc tests are useful for sorting comparisons, in a case like this – where a tendency seems to be consonant with reasoned expectations despite a lack of statistical significance – it is important to consider the possibility of a 'true' finding. This assortment of data is consistent with the characterisation of two different poles of involvement, the axis upon which much of the data differed in the work of Biber, Chafe, and Jang.

For **downtoners**, a significant difference between the registers did not obtain ($F = 2.781$, $p = .905$ n.s.). Due to the lack of interpretable comparisons, this feature will not be discussed again.

Commentary. While the statistical power was low in these comparisons, they show a clear tendency for sports reportage and radio interview to be differentiated from news scripts and, in terms of hedges, traditional narrative. Radio interview showed high levels of amplifiers, emphatics and hedges, while sports reportage had the highest mean for amplifiers and one of the highest for hedges. This result is generally consistent with previous research; radio interview is a classically involved register – bound up with interaction and opining – while the news script register is primarily factual and detached.

One of the characteristics that makes radio interview compelling to its audience is its subjective, personal element, where different positions are expressed by, and contrasted between, different participants. In radio interview, amplifiers are used in various ways, but they often function to emphasise some quality central to a position: to give it more communicative weight so that it will be perceived as important. Hedges are seen to: indicate inexactitude; suggest alternatives or hypotheses; and downplay an assertion. Emphatics are used to convey a general sense of conviction towards a proposition or indicate agreement with a prior utterance, such as from another participant. All of these functions are characteristic of involved discourse, where the expression of personal opinions and interaction between participants has a more central role than the exchange of information.

In sports reportage, amplifiers are used as mentioned above, but even more often to indicate the physical positions of referents on field (players, the ball, etc.). Hedges are used mainly in the first way mentioned above – to indicate that references to time or physicality (movement, location) are inexact or approximate. Consequently, the use of these features is not primarily for the expression of personal stance and other functions related to interaction³⁹ but, instead, for extra-textual references to space and time.

³⁹ But this aspect does appear, especially during inter-play commentary.

Some examples follow illustrating these features:

Amplifier

- (91) *thathar a' cumail a-mach gu bheil a' phàpa (sic.) a th' againn an-dràsta gu bheil e gu math teann gu math cruaidh mar a chanas iad ro right-wing (INTER)*
 'it is being maintained that the pope we have right now that he is **very** restrictive very hard as they say **too** right-wing'
- (92) *Nichol air taobh a-staigh leth Yugoslavia ga feuchainn suas na loidhne ach bha siud ro fhad air falbh bho Johnson agus chaidh e a-mach (SPORT)*
 'Nichol inside Yugoslavia's side trying [to put] it up the line but that was **too** far away from Johnson and it went out'

Hedge

- (93) *mar a bha cuideigin tha mi a' smaoinichadh a's an riaghaltas againn fhìn a' ràdh gu bheil a cheart uimhir a chead neo a chòir aig Breatainn 's dòcha am bomb a thilgeil air New York ri linn 's gu bheil ceannaircich bhon IRA a' fuireach an-sin (INTER)*
 'as somebody I think in our own government was saying that Britain has just as much justification **perhaps** to drop the bomb on New York because terrorists from the IRA are living there'
- (94) *feuchaidh iad buille 's dòcha bho dà fhichead slat no mar sin (SPORTS)*
 'they will try a kick from **perhaps** twenty yards or about that'

Emphatic

- (95) *<a> a Dhòmhnaill ma thèid mi ugaibh fhèin an toiseach ma-thà a bheil sibhse draghail man a bheil a' tachairt?*
* thà draghail dha-rìribh tha seo a' bualadh air daoine neo-chiontach saoilidh mise (INTER)*
<a> Donald if I may go to you first then are you bothered about what's happening?
* yes bothered indeed this is affecting innocent people I think*

In example (91), the speaker uses the amplifiers to emphasise certain characteristics about the pope, as well as saying that he possesses an *excessive* level of a certain quality – being 'right-wing'. While the use of amplifiers in this case related to a particular argumentative position, the use of the amplifier in example (92) simply relates to the physical positions of a ball and player in a football game. Example (93) uses a hedge, again in the context of an argumentative position. The logic of this example goes that, if America can bomb a country because it is harbouring terrorists, what would be wrong with Britain bombing New York, since terrorists from the IRA are known to be in residence there. The hedge is used to downplay conviction to this incendiary statement, effectively saying that maybe it follows as a logical argument, but not necessarily. On the other hand, the hedge in example (94) only serves to indicate that the measurement provided in the utterance is inexact. Finally, in example (95), speaker 1 asks a question of speaker 2 to which he responds with an affirmative response along with an emphatic. The emphatic operates to convey an increased conviction in the statement; not only is he bothered, he is very bothered about what is happening.

6.5.2 Verb Types

Three verb classes were investigated in this section: private; public; and suasive. Biber (1988) defines these as follows: **private verbs** involve mental states (*creidsinn* 'believing') or "nonobservable intellectual acts" (p. 242) such as *mothaich* 'notice'; **public verbs** denote actions which can be publicly witnessed, such as *ràdh* 'say' or *innse* 'tell';

suasive verbs are those that "imply intentions to bring about some change in the future" (p. 242) such as *iarr air* 'ask for' or *ionnsachadh* 'teach' (can also mean 'learn'). Chafe (1982) and Biber (1988) found that private verbs signify involvement while Jang (1998) reports them as indicating personal expression of emotion. Public verbs were found by Biber (1988) to be frequent in narrative and Jang (1998) found them common in interactive registers. Suasive verbs were found by Biber (1988) to be one of the main features of persuasive registers, such as editorials.

Data and commentary. The means of the different types are presented below:

Table 80: Results: verb types

Register	Private	Public	Suasive
Conversation	8.89	5.97	.21
Radio Interview	13.43	6.34	.34
Sports Reportage	5.63	3.03	.30
Traditional Narrative	8.52	7.30	1.41
SPEECH	8.89	5.62	.58
Fiction	10.83	10.09	.72
Formal Prose	3.04	4.35	.32
News Scripts	4.77	15.58	2.53
Popular Writing	8.26	9.54	.81
WRITING	6.65	10.02	1.13
Total	7.77	7.82	.85

For private verbs, there was a significant overall difference between the registers ($F=5.059, p < .001$). Significant pairwise comparisons were as follows:

Table 81: Sorted data: private verbs

13.43	●	Radio Interview
10.83	●	Fiction
8.89	●	Conversation
8.52	●	Traditional Narrative
8.26	●	Popular Writing
5.63	●	Sports Reportage
4.77	●	News Scripts
3.04	●●●●	Formal Prose

Radio interview had the highest mean for the feature, followed by fiction. Formal prose, in contrast, had the lowest mean, being less than one quarter that found in radio interview. There were significant pairwise comparisons between formal prose and the four highest scoring registers, of which three are spoken. This result indicates that private verbs are generally more common in spoken, involved registers such as radio interview and conversation, but also present in high levels in narrative. Formal prose is the epitome of a detached register and it is not surprising that it had a very low return for the feature. Although it did not differ significantly from any other register, it is notable that news scripts had the second lowest return for private verbs. Some examples follow from the four highest scoring registers:

- (96) [a' bruidhinn air daoine ann an tug of war] uell nas mò a tha iad nas brùideil a tha iad
 tha mi a' creidsinn sin as fheàrr gun teagamh (INTER)
 [speaking about participants in tug of war] 'well the bigger they are the tougher
 they are I believe'

- (97) *Agus on a thàinig an deich millean not [sic] gu telebhisean Gàidhlig bhiodh tu às do chiall mura do smaoinich thu air sin.* (FICT)
'And since the ten million pounds came to Gaelic television you would have to be out of your mind unless you **thought** of that'
- (98) *uell na Dòmhnallaich sin bha iad càirdeach do Sheonaidh tha mi a' smaointinn math dh'fhoadte* (CONV)
'well those MacDonalds they were related to Johnny I am **thinking** perhaps'
- (99) *feumaidh mi fhèin falbh dha h-iarraidh agus an gille còmhla rium thoradh chan earb sinn ri duine sam bith i* (TRAD)
'I must go to get it and the boy with me because we won't **trust** it with any other person'

Public verbs also differed significantly between the registers ($F = 5.924$, $p < .001$). The following significant pairwise comparisons emerged:

Table 82: Sorted data: public verbs

15.58	●●●●●	News Scripts
10.09	●●●●●	Fiction
9.54	●●●●●	Popular Writing
7.30	●●●●●	Traditional Narrative
6.34	●●●●●	Radio Interview
5.97	●●●●●	Conversation
4.35	●●●●●	Formal Prose
3.03	●●●●●	Sports Reportage

News scripts clearly had the highest mean and was significantly different from most of the other registers, as indicated by the arrows and illustrated by the shading. Popular writing also differed significantly from sports reportage, which had the lowest mean for public verbs, seconded by formal prose.

With its emphasis on reporting public events and the speech of public figures, it is consistent that news scripts would present the highest count of public verbs. Sports reportage, on the other hand, is virtually devoid of interaction, except in the short commentaries between plays. The results differ somewhat from Biber (1988) who found that narrative had the highest return for private verbs. However, as can be seen from the Table 82 above, fiction had the second highest level overall while traditional narrative had the highest for a spoken register. On the other hand, the results are strikingly different from Jang (1998) who stated that public verbs were primarily associated with interactive registers.

Examples (100) and (101) are from news scripts of public verbs:

- (100) *Tha eòlaichean ag ràdh gur e droch togalaichean [sic] nach seasadh ri crith-thalmhainn as coireach gun do bhàsaich uidhir.*
'Scientists are **saying** that bad houses which could not stand up to an earthquake are the reason that so many died.'
- (101) *Tha buidheann na Slàinte is Sàbhailteachd dol a dh'fhoillseachadh aithisg a chàineadh Railtrack a thaobh stait loidhnichean iarainn Bhreatainn.*
'The Health and Safety Department is going to **publish** a report that would criticise Railtrack concerning the state of Britain's railways.'

For **suasive verbs**, the Kruskal-Wallis test showed a significant overall difference ($X^2 = 28.442$, $p < .001$). News scripts evinced the most, as can be seen in the table of means above (Table 80), while conversation evinced the least. If an example of a persuasive,

editorial register had been included, we would have expected to find it ranked higher than news scripts. While the news script register is not, itself, an example of this kind of register, many of the quotes contained within the scripts are from individuals who are trying to influence the opinions of others or cause some change in the listener. In this sense, one may say that there is a sub-register present within it. Some examples follow from news scripts:

- (102) *Coinnichidh iasgairean chlàibh a' Chost an Iar a-màireach, agus iad **ag iarraidh** air a' Riaghaltas a-rithist casg a chur air tràlairean bho na lochan mara.*
'West coast creel fisherman will meet tomorrow, **wanting** the government again to put a ban on trawlers from the sea lochs.'
- (103) *Chaidh a **mhòladh** le luchd-sgrùdaidh an riaghaltais gum faodadh sgudal niuclasach na h-Alba aig ìre ìosal a stòradh aig ionad Dhùnrrath.*
'It was **suggested** by government researchers that nuclear waste could be stored to some extent at the Dounreay centre'
- (104) *Tha poilis Alba air iomairt ùr a chur air chois, son feuchainn ri bhith a' **toirt air** draibhear na criosan sàbhalaich aca chleachdadh.*
'Scottish police officers have begun a new campaign, trying to **make** drivers use their safety belts.'

6.5.3 English Borrowing

Prevalence of English borrowing has been taken by several researchers (see Chapter II) as indicating lowered formality. The borrowing data were split into two groups: nouns and non-nouns. Within the nouns, two features were counted: nouns assimilated to Gaelic orthography/phonology and nouns that were left unassimilated. An example of the first is *copadh* for 'coping'. Here, the stem of the English word ('cop(e)-') has been imported with the addition of the Gaelic suffix used with verbal nouns. Unassimilated words were left in their normal English form. For non-nouns, only unassimilated totals were counted, as assimilated totals did not reach high levels.

Data. Table 83 presents the overall means for these three features:

Table 83: Results: English borrowing

Register	Assimilated noun	Unassimilated noun	Unassimilated non-noun
Conversation	2.95	14.76	5.46
Radio Interview	1.41	6.27	1.81
Sports Reportage	.48	.49	.09
Traditional Narrative	.60	.22	.00
SPEECH	1.36	5.39	1.84
Fiction	1.57	2.17	.12
Formal Prose	3.36	.88	.12
News Scripts	3.14	2.66	.15
Popular Writing	1.89	2.09	.55
WRITING	2.50	1.97	.25
Total	1.93	3.68	1.05

The occurrence of **assimilated nouns** was examined with the Kruskal-Wallis test and found to differ significantly in the registers ($X^2 = 14.274$, $p = .047$). Formal prose had the most, as seen below, and the register with the least was sports reportage. The sorted data follow for comparative purposes with examples from the texts provided:

Table 84: Assimilated English nouns

3.36	Formal Prose
3.14	News Scripts
2.95	Conversation
1.89	Popular Writing
1.57	Fiction
1.41	Radio Interview
.60	Traditional Narrative
.48	Sports Reportage

Examples:

CONV 'sa bhunkhouse; *brainseachan* '(bank) branches'; *restadh* 'resting'

INTER *treanadh* 'training' (repeats 4 times)

SPORT *air a'* phost; *air a'* chórner

TRAD *trungaichean* 'trunks'

FICT *suimeant* 'cement'; *glosaidh* 'glossy (magazine)'; 'sa chontract

FORM na *Gypsidhean* 'the Gypsies'; *ailealan* 'alleles'

NEWS *saidheans* 'science'; *poileasaidh* 'policy'; *democrasaidh* 'democracy'; *helicoptearan*

POP an *Leadaidh* 'the lady'; *sistam* 'system'; an *TBh* 'the TV'

As can be seen, the extent of assimilation ranges from simple application of lenition to fully gaelicised structures such as *brainseachan*, with phonological modification (English [t] → Scottish Gaelic [ʃ]) and added morphemes (plural *-eachan*).

For unassimilated nouns, a significant difference also obtained ($X^2 = 38.652$, $p < .001$). There was a clear contrast in the spoken registers between conversation, at 14.76 instances per 1000 words, and traditional narrative, at .22 instances (cf. formal prose, at .88 instances). The sorted data and examples follow:

Examples:

CONV forecast; turkey burger; gym; lemon sole; horse box

INTER *am* medication; coach, exercise, elephant

SPORT *aig a'* pole; *dhen a'* game; *gu* corner *a'* *bhoesa* 'to the corner of the box'

TRAD chief; lid (occurs as a gloss of *còmhlà*)

FICT concert; contrast; deal; bullshit

FORM an Arctic; epic; Sanskrit; introduction

NEWS tribunal; hepatitis C; e-mail; cervical cancer

POP grant; wireless; family

Table 85: Unassimilated English nouns

14.76	Conversation
6.27	Radio Interview
2.66	News Scripts
2.17	Fiction
2.09	Popular Writing
.88	Formal Prose
.49	Sports Reportage
.22	Traditional Narrative

Most of these are self explanatory. The first two in the sports register row are in contrast to the ones cited above, which evince lenition as they are definite and in the dative case.

Finally, for **unassimilated non-nouns**, there was also a significant contrast between the registers ($X^2 = 27.072$, $p < .001$). Again, the most extreme contrast was between conversation, at 5.46 instances, and traditional narrative, which did not evince a single instance in all of the texts. The sorted data and examples follow below:

Table 86: Unassimilated English non-nouns

5.46	Conversation
1.81	Radio Interview
.55	Popular Writing
.15	News Scripts
.12	Fiction
.12	Formal Prose
.09	Sports Reportage
.00	Traditional Narrative

Examples:

CONV	fifty-five; still <i>ann</i> 'still there'; supposed; brainy
INTER	important; really; eighteen twenty; right wing; alright
SPORT	<i>airson</i> spell 'for a spell [of time]'
TRAD	N/A
FICT	<i>gabh</i> straight <i>a-steach</i> 'go straight in'
FORM	record (vb.)
NEWS	aluminium (adj.); Chinook (adj., modifying 'helicopter')
POP	fussy; knackered; extended

Commentary. In order to understand borrowing in each of the registers, it is necessary to build a picture incorporating the three different features. If formal prose is considered, it is seen to have the highest level of assimilated English nouns, but very low levels of unassimilated nouns and non-nouns. This pattern indicates that authors of formal prose eschew conspicuous English loans, preferring to present them in gaelicised form. On the other hand, in conversation, there is a high level of assimilated English nouns, but an even higher level of unassimilated ones and unassimilated non-nouns. Consequently, it appears that conversation contains many loans regardless of form. This finding agrees with the observation of other researchers – such as Macaulay (1982a), Ni Laoire (1988) and Dorian (1994a) – that colloquial usage correlates with greater occurrence of borrowing. When comparing conversation to radio interview, however, one sees less borrowing altogether. This is ostensibly indicative of the greater self-consciousness emerging in individuals who speak on the radio: they are less inclined to present an English form. For sports reportage, there were also very few loans, due to the effort on the part of the BBC Gaelic department to establish terminology for football coverage in the language (see §4.1.3.3 for examples). Moreover, the limited semantic range of the register enables this effort to be comprehensive. In contrast to all of these cases is traditional narrative. The results show that this register has low levels of loans altogether. This is a product of at least two factors: traditional narrative is a formal register, with most *seanchaidhean* 'tradition bearers' investing their stories with inherent cultural worth; and the context of such stories is that of a world pre-dating pervasive English influence – there are few loans because there is little need for them.

6.5.4 Numerals

The author could find no references for the register variation of numerals in the literature (save for the English grammar of Biber *et al.* 1999). In Chafe's terminology, use of numerals is an example of integration – the packaging of information in a clause – similar

to other kinds of nominal modification. In the present study, no distinction was made between cardinal and ordinal numbers, nor those used pronominally *versus* those used attributively (i.e. 'the three over there' *vs.* 'the three people over there'). In Biber *et al.* (1999), news was found to have the greatest level of numerals overall followed by academic prose.

Data and commentary. There was a significant difference between the registers for numerals ($F = 5.059, p < .001$). The means obtained are as follows:

Table 87: Results: numerals

Register	Mean
Conversation	6.77
Radio Interview	13.77
Sports Reportage	7.32
Traditional Narrative	6.36
SPEECH	8.28
Fiction	6.80
Formal Prose	18.68
News Scripts	22.38
Popular Writing	10.61
WRITING	14.82
Total	11.55

A number of significant pairwise comparisons also emerged:

Table 88: Sorted data: numerals

22.38	●●●●●	News Scripts
18.68	●	Formal Prose
13.77	●	Radio Interview
10.61	●	Popular Writing
7.32	●	Sports Reportage
6.80	●	Fiction
6.77	●	Conversation
6.36	●	Traditional Narrative

The results are very close to those found by Biber *et al.* (1999). News scripts had the highest occurrence level for numerals and, as indicated by the arrows, differed significantly from five of the registers in the corpus: traditional narrative; conversation; fiction; sports reportage; and popular writing. Formal prose had the second highest mean for numerals, but did not differ significantly from any of the other registers in the corpus. However, based upon the position of these two registers, it is reasonable to suggest that the occurrence of numerals correlates positively with informativity.

6.6 Lexical Specificity

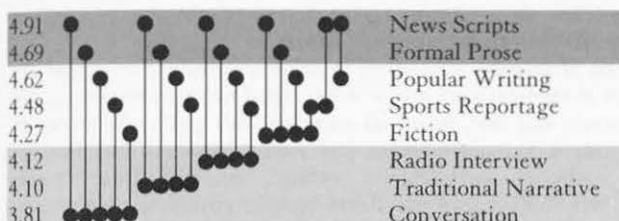
Two measures of lexical specificity were used: the average number of letters per word and the type/token ratio, which measures the number of different words in a text. Biber (1988) found greater mean word length correlated positively with informativity in registers. In terms of the type/token ratio, studies have found that it is higher in writing (Drieman 1962; Chafe and Danielewicz 1987) and that it is also indicative of an informational orientation (Biber 1988; Jang 1998), and particularly marked in news (Biber *et al.* 1999).

Table 89: Results: lexical specificity

Register	Characters per word	Type/token ratio
Conversation	3.81	48.45
Radio Interview	4.12	46.53
Sports Reportage	4.48	40.73
Traditional Narrative	4.10	43.72
SPEECH	4.13	44.77
Fiction	4.27	53.56
Formal Prose	4.69	50.14
News Scripts	4.91	56.25
Popular Writing	4.62	52.18
WRITING	4.64	53.06
Total	4.38	48.91

Data and commentary. There was a significant difference between the groups for **average characters per word** ($F = 44.454$, $p < .001$). Significant pairwise comparisons obtained between most of the registers:

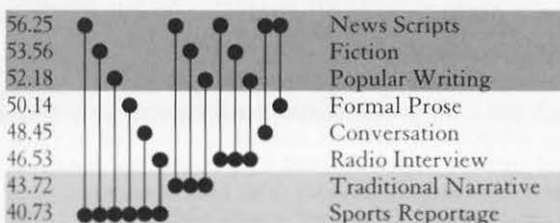
Table 90: Sorted data: average characters per word



It seems that all of the written registers save fiction returned relatively high counts for average characters per word. News scripts had the highest count by far, followed closely by formal prose and popular writing. At the other end of the spectrum, conversation was markedly different from most of the other registers. The return for sports reportage may be a bit inflated due to the length of some of the surnames encountered, which occur with a high frequency. This finding is consistent with Biber (1988), indicating that more informational, literate registers in Scottish Gaelic tend to have longer words.

A significant difference also obtained for the **type/token ratio** ($F = 24.99$, $p < .001$). The significant pairwise comparisons were as follows:

Table 91: Sorted data: type/token ratio



Here, news scripts again has the highest return, followed by fiction and popular writing. At the other end of the scale are sports reportage and traditional narrative. A statistically significant contrast occurred between sports reportage and every other register save traditional narrative. News scripts also differed significantly from five out of the 7 other registers.

In view of these two results, we see that news scripts both have the most extensive lexicon and use the longest words. This reflects the great variety of referents in the register as well as its high concentration of nouns. Fiction has a varied lexicon, but tends to use shorter words than one half of the other registers. Formal prose, on the other hand, tends to use longer words, but does not have a particularly varied lexicon. This may be due – as Biber *et al.* (1999) mention about academic prose – to the greater repetition of especially technical terms in this register. Conversation tends to use the smallest words, but has the most varied lexicon of the spoken registers. Sports reportage has the narrowest lexical range of the corpus. This is likely to be a factor of the very constrained production circumstances of this register. Overall, as has been found in studies on other languages, written registers in Scottish Gaelic have a more varied vocabulary than spoken registers and also tend to use longer words.

6.7 Summary of Chapter Six

This chapter surveyed a wide range of morphosyntactic features that were suspected, based upon prior research and observation, to be relevant to register variation in Scottish Gaelic. A number of these features marked strong differences between the registers. The following discussion summarises these differences in terms of the most apparent functional or contextual correlates associated with the features.

The largest functional class of features is involved in differentiating registers that are primarily informational from those which are not. Passives, total nouns, complex nouns, genitives, prepositions, numerals, and long words all occurred with greater frequency in the more informational registers, i.e. news scripts and formal prose. These features function to: emphasise actions and objects whilst de-emphasising actors (passives/impersonals); pack information into the clause (nouns, genitives, prepositions, and numerals); and increase specificity of reference (long words). Additionally, clausal negation, which is linked to more imaginative or involved discourse, had a low occurrence level in these texts.

Other features were associated with a contrast between narrative and non-narrative discourse: present tense; past tense; indefinite 2; and pronouns. The first of these – present tense – was found to be relatively scarce in narrative while the following three were much more common in narrative than other types of registers. Past tense is the classic marker of narrative discourse, but the indefinite 2 can also refer to past time frames. Finally, pronouns seem to be more common in narrative due to the heavy need for referent tracking.

Interactive discourse was distinguished by several features: interrogatives; imperatives; indefinite marking; and private verbs. The first two of these are superficial correlates of dialogue. Although interrogatives were clearly more common in conversation and radio interview, imperatives were most prominent in traditional narrative. This was explained as being due to the setting in which the conversations took place; had there been more procedural discourse – such as that surrounding communal work – or had all the conversations taken place over the dinner table, perhaps there would have been more imperatives found. Also, the 'rough and ready' nature of traditional narrative may select more impolite strategies for giving orders, while these are somewhat proscribed by social conventions in most conversational contexts. Indefinite marking increases the vagueness of a reference and tends to indicate a more fragmented style. This kind of construction is interactive insofar as it places an onus upon the addressee(s) to provide a reference for

themselves. Private verbs communicate the opinions, perceptions, and feelings of participants and were most common as a whole in the two interactive registers, conversation and radio interview.

Differences in production constraints were associated with the distribution of a number of features: attributive adjectives; complex noun phrases; grammatical concordance; and unassimilated English loans. In general, registers with production time restrictions (i.e. conversation and radio interview) tend to have a simpler syntax and morphology than those permitting time for editing (i.e. all written registers except for news scripts). However, self-consciousness – even in spontaneous spoken registers – can have an impact on linguistic form, as was found in the difference between conversation and radio interview regarding unassimilated English loans: conversation had more.

Finally, the distribution of demonstrative adjectives – most frequent in formal prose and conversation and least frequent in news scripts and sports reportage – indicated a sensitivity to discourse freedom. While news scripts and sports reportage are bound to maintain a continuous flow of information, formal prose and conversation are free to pick up previous points of reference. Demonstrative adjectives are one of the devices whereby referents can be reactivated.

7 Noun Phrase Grammar and Complexity

This chapter investigates in more depth two features covered in Chapter 6 which were found to robustly diverge across the different registers: concordance of case marking and noun phrase complexity. A frequently encountered position concerning Scottish Gaelic register variation is that formal uses of the language tend to correlate with the appearance of more conservative nominal forms. The Goidelic languages are notable for the complex ways in which case, gender, and number interact in their nominal morphology. (A sketch of the system is presented in §2.2 of Appendix 1.) As in many other languages, this system has been in a process of gradual simplification for as long as the written record is extant. Today, it is clear that there is a wide range of variation in the use of nominal forms, but there has been no attempt, thus far, to quantify those most subject to variation, nor to determine which registers evince the most conservative usage. The first sub-chapter examines these topics, breaking the data down into three cases – nominative, dative, and genitive – as well as adjectival usage. Furthermore, by collapsing the corpus data together as a whole, it investigates which combinations of case, gender, and number tend to produce the most cases of nonconcordance overall.

In the second sub-chapter, the differences in noun phrase complexity are investigated between two different discourse types: spontaneous conversation and formal prose. Previous work, such as Chafe (1982), Biber (1988), and Miller and Weinert (1998), has indicated that speech and writing differ in the amount of information that they pack into clauses and phrases. Writing tends to be more informationally compressed, as seen in its greater proportion of modified noun phrases. This result was replicated in the present study in §6.2.3, where there was a solid difference between formal prose and conversation in their concentrations of complex noun phrase. Formal prose was found to have 15.83 complex NPs per 1000 words while conversation had only 2.36. This chapter investigates in finer detail the ways in which these two registers differ. It examines the differences in the syntactic functions of complex NPs in the two registers (e.g. as subject or object of the clause) as well as the differences in the frequency of the constituents used in their respective complex noun phrases (i.e. attributive adjectives, relative clauses, prepositional phrases, and genitives). The results indicate that the complex NPs of formal prose tend to show greater lexical range, have a greater number of modifying NPs which are themselves modified, and show more 'upmarket' modifiers. In addition, a greater proportion of complex NPs in formal prose occur as syntactically integrated constituents (e.g. the subject of a transitive clause) while more occur in extraclausal positions in conversation.

7.1 Case Marking in Scottish Gaelic Registers: Concordance and Nonconcordance

The statistical analysis in §6.2.4.2 revealed some interesting differences in case marking between the registers. To recapitulate, sports reportage and news scripts were seen to have had the highest proportion of nonconcordant marking while fiction, traditional narrative and formal prose had the lowest. Before carrying on from this finding, it seemed important to establish whether there was an interaction between the number of nouns that a text exhibits and its level of grammatical nonconcordance; in other words, whether a register that is very noun rich (e.g. news scripts) has more of a tendency towards nonconcordance than one that is relatively noun poor (e.g. conversation). To investigate this possibility, Kendall's tau-b nonparametric test of correlation was used to measure the relationship between two variables: the overall count of nouns; and the ratio of nonconcordant items to concordant items. The test revealed that there was not a significant relationship between the two features ($r = .126, p = .155$ n.s.). Therefore, in this corpus, there is not a tendency for texts to evince higher levels of noun nonconcordance the more noun rich that they are.

Using *LinguaStat*, all instances of noun nonconcordance in the corpus were located, with cases separated for each register. Similarly, another search was done for all instances of noun concordance. The resulting data are displayed in the tables that follow, organised per case category: nominative; dative; and genitive. Broad columns are organised per number/definiteness and divided into 'C' concordance and 'N' nonconcordance. The nonconcordance columns have been shaded for ease of reference.

7.1.1 Nominative Case

As seen below, there were very few cases of nonconcordance in the nominative. In the spoken texts, there are examples such as *an fhitheach* (~*am fitheach* m.; traditional narrative), *an ùrachadh* (~*an t-ùrachadh* m.; sports reportage); *bha an cuid cluich aig McClair* (~*a' chuid* f.; sports reportage), and *a' phàpa a th' againn* (~*am pàpa* m.; radio interview). In the written texts we find *an comhairle* (~*a' chomhairle* f.; news scripts), *an ionad* (~*an t-ionad* m.; formal prose), *'s e am faireachdainn a bha sinn ...* (~*an fhaireachdainn* f.; popular writing). In nearly all cases, the nonconcordance is due to the lenition of masculine nouns or the non-lenition of feminine nouns. Following from this, most examples are definite. One interesting case of an nonconcordant indefinite noun is a hypercorrection in a news script: *eadar bhuidheannan* (~*buidhnean*). The preposition *eadar* is usually taken to govern nominative case rather than the genitive, as it does here.

Table 92: Noun concordance data: nominative case, feminine

No./Definiteness Concordance	Pl. Ind.		Sing. Def.		Sing. Ind.	
	C	N	C	N	C	N
Conversation	1.11		4.82		10.51	
Radio Interview	1.80		5.88		10.34	
Sports Reportage	1.35		3.37	0.10	12.03	
Trad Narrative	0.29		9.57		10.05	
Fiction	2.31		11.36		19.15	
Formal Prose	1.57		6.59		15.45	
News Scripts	2.68	0.10	6.64	0.30	27.85	0.10
Popular Writing	2.46		9.65	0.10	20.23	
Average	1.70	0.10	7.24	0.17	15.70	0.10

Overall, there are fewer instances of nonconcordant feminine nouns than masculine nouns. This may be partly due to the fact that feminine nouns are not as prevalent as masculine nouns in general, which can be seen by comparing the column averages at the bottom of each table.

Table 93: Noun concordance data: nominative case, masculine

No./Definiteness Concordance	Pl. Def.		Sing. Def.		Sing. Ind.	
	C	N	C	N	C	N
Conversation	1.42		11.30		24.26	
Radio Interview	4.55		9.70	0.19	28.65	
Sports Reportage	2.12	0.10	11.35	0.19	16.07	
Trad Narrative	3.51		21.77	0.10	26.45	
Fiction	1.79		17.89		32.72	
Formal Prose	4.92		13.38		31.39	0.10
News Scripts	3.50	0.10	10.01	0.69	38.85	0.50
Popular Writing	3.90		11.30	0.10	34.92	
Average	3.21	0.10	13.34	0.25	29.16	0.30

The highest concentration of nonconcordance overall is clearly in the news scripts. To measure its disparity with the other registers, one can compare the levels of concordant nouns in the news scripts with the others and then compare these figures with the levels of nonconcordance. Doing this, one finds that news scripts evince 10.01 concordant singular, definite, masculine nouns per 1000 words with 0.69 that are nonconcordant. While its level of nonconcordance in this category is the highest of the registers, its level of concordance is the second lowest (cf. the levels of C and N for traditional narrative). This tells us that news scripts did not have a preponderance of these types of nouns compared to the other registers, but more often presented a nonconcordant form.

7.1.2 Dative Case

The dative (or prepositional case) is the name given to the case assigned by most prepositions in Scottish Gaelic. Generally speaking, nominative forms are used as the basis for the dative, but suppletion does occur with a limited set of irregular nouns, e.g. *bò* nom. → *air a' bhoìn* dat. Definite masculine nouns in the dative are either lenited – if they begin with a lenitable noun (e.g. *a's a' chàr*) – or receive the prefix *t-* if they begin with *s* or *s* + sonorant [l, n, r], e.g. *leis an t-snàmh*. Definite feminine nouns receive the same treatment as masculine ones, but they also indicate the dative by palatalisation of terminal consonants and/or vowel changes when possible. Depending on the preposition implicated, indefinite nouns may or may not be lenited. However, in conservative usage, all indefinite dative feminine nouns are palatalised, with vowel changes, if their morphology permits it. Diagrams illustrating these features are available in §2.2.1.4 of Appendix 1.

A quick look at the tables below indicates many more nonconcordant forms than the nominative did. Beginning with the spoken registers, we find the following: *an dà chailleach*⁴⁰ (~*chaillich* f.; conversation); *gann de chlach* (~*chloich* f.; conversation); *comhla ris a' chlànn* (~*chloinn* f.; radio interview); *le a chas* (~*chois* f.; sports reportage); *a's a' bhucais* (*bhucais* m.; sports reportage). Some examples from the written registers follow: *ri a nighean* (~*nighinn* f.; fiction); *bhon a' bhuidheann* (~*bhuidhinn* f.; news scripts); *thug air an criubha an cùrsa aca atharrachadh* (~*air a' chriubha* m.; news scripts). As is clear, most of the nonconcordant forms stem from the lack of palatalisation in dative feminine nouns. In the nonconcordant masculine nouns, which are less frequent, nonconcordance is due to feminised hypercorrection (applying palatalisation when it should not occur); failure to lenite; and misplacement of '*t-*' (e.g. *aig an aon t-sagart* ~ *aig an aon s(h)agart*, cf. *aig an t-sagart*; popular writing).

Table 94: Noun concordance data: dative case, feminine

No./Definiteness	Sing. Def.		Sing. Ind.	
	C	N	C	N
Concordance				
Conversation	3.71	0.24	2.45	0.16
Radio Interview	5.79	0.38	4.93	0.38
Sports Reportage	6.16	0.67	1.73	0.48
Trad Narrative	8.69		4.98	0.20
Fiction	10.94	0.21	11.05	0.32
Formal Prose	13.18		14.27	0.10
News Scripts	12.98	0.50	14.47	0.89
Popular Writing	14.28	0.21	10.68	0.10
Average	9.47	0.37	8.07	0.33

⁴⁰ Technically, this is an instance of dual number, but its formation in feminine nouns is identical to that of the dative.

The highest level of nonconcordance for dative feminine nouns that are singular and definite appears in sports reportage, seconded by news scripts. For indefinite forms, these two registers are reversed. Formal prose appears to be the most conservative register for feminine dative forms as it has the highest incidence of concordant forms and the lowest incidence of nonconcordant forms.

Table 95: Noun concordance data: dative case, masculine

No./Definiteness	Pl. Ind.		Sing. Def.		Sing. Ind.	
	C	N	C	N	C	N
Concordance	0.55		5.53	0.08	5.85	
Conversation	3.23		7.30	0.38	9.39	
Radio Interview	0.67		14.53	0.10	10.10	
Sports Reportage	0.39	0.10	16.69	0.10	7.91	
Trad Narrative	2.52		15.89		16.83	
Fiction	4.82		16.43		19.48	
Formal Prose	4.66	0.10	8.92	0.40	15.86	0.20
News Scripts	4.42		14.38	0.10	17.46	
Popular Writing	2.66	0.10	12.46	0.19	12.86	0.20
Average						

For masculine dative forms, the greatest frequency of nonconcordance occurred with nouns that were singular and definite. Again, news scripts had the highest level of nonconcordance overall.

7.1.3 Genitive Case

As is easily grasped from the tables below, the genitive presents the greatest overall levels of nonconcordance, while the actual levels of use are much lower than the other cases (besides the vocative, which was very rare in the corpus). Plural genitives are characterised by lenition when indefinite and the article *nam/nan* when definite. In addition, especially with Type 1 masculine nouns (see §2.2.1 in the Grammar), there can be a reversion back to the nominative singular form in the plural genitive. Examples of nonconcordance in the spoken registers for genitive plural nouns follow: *sgath giomaich* (~*ghiomach* m.; conversation); *beagan chaoraich* (~*chaorach* f. irreg.; conversation); *a-steach a leth na h-Astràilianach* (~*nan Astràilianach* m.; sports reportage); *air feadh na h-eilean* (~*nan eilean* m.; radio interview); *aig ceann cùisean* (~*chùisean* f.; radio interview). For the written registers, we find: *a' togail dealbhan* (~*dhealbhan* m./f.; popular writing); *airson na dùthchannan* (~*nan*; news scripts); *a' coimhead builgeanan* (~*bhuilgeanan* m.; fiction).

The nonconcordant singular genitives are much more frequent than the plural ones. In the spoken registers, one locates the following examples: *mu dheidhinn fear no tè* (~*fir* m.; conversation); *a' dèanamh crìoch* (~*crìche* f.; traditional narrative); *a' tuigsinn an obair* (~*na h-obrach* f.; radio interview); *airson a' bhall* (~*a' bhuill* m.; sports reportage). The nonconcordance is due to lack of genitive definite articles and lack of the specific genitive forms, many of which are irregular, for both masculine and feminine nouns. In the written registers, the same kinds of nonconcordance occur: *am measg buidheann* (~*buidhinn* f.; news scripts); *ag iarraidh dearbhadh* (~*dearbhaidh* m.; formal prose); *a' cur an dreach fhèin* (~*dreacha* m.; popular writing); *far a' bheing* (~*na beinge* f.; fiction).

In one or two registers, nonconcordant items are distributed across a very small number of lexical items, boosting the level of nonconcordance out of proportion. For example, in light of the high score of sports reportage for nonconcordant singular, definite nouns below, the main words encountered are ones that are frequently repeated in that register, almost to the point of being 'register markers': *bròg* 'shoe', *loidhne* 'line', *cluich* 'play'. However, in other registers, such as news scripts the items are more disparate.

To ease comparison making in the tables that follow, column maximums are indicated in black colour.

Table 96: Noun concordance data: genitive case, feminine

No./Definiteness	Pl. Def.		Pl. Indef.		Sing. Def.		Sing. Ind.	
	C	N	C	N	C	N	C	N
Concordance								
Conversation	0.24	0.08	0.55	0.08	2.29	0.16	1.26	0.40
Radio Interview	0.38	0.28	0.66	0.38	2.66	0.76	2.85	0.95
Sports Reportage	0.19	0.10	0.29	0.10	4.62	1.73	3.46	0.48
Trad Narrative	0.88		0.49	0.10	2.93	0.20	2.15	0.39
Fiction	0.95		0.84		4.63	0.21	4.52	0.74
Formal Prose	0.89		0.59		6.30	0.39	6.40	0.30
News Scripts	0.89	0.30	1.09	0.59	8.52	0.59	8.42	2.28
Popular Writing	1.23		0.92	0.10	9.24	0.31	7.70	1.13
Average	0.71	0.19	0.68	0.23	5.15	0.54	4.60	0.83

For feminine nouns in the genitive, there seems to be a tendency for news scripts to have the highest level of nonconcordance, with the next highest level – in 2 out of 4 cases – in radio interview. This is not surprising as both are affected by production constraints. Formal prose is consistently among the most conservative registers, along with fiction. Although conversation has very few nonconcordant feminine genitives which are singular and definite, it also has very few concordant ones, as seen above.

Table 97: Noun concordance data: genitive case, masculine

No./Definiteness	Pl. Def.		Pl. Ind.		Sing. Def.		Sing. Ind.	
	C	N	C	N	C	N	C	N
Concordance								
Conversation	0.71		0.47	0.40	4.27	0.40	1.90	0.87
Radio Interview	0.85	1.14	0.76	0.57	3.32	0.28	3.41	1.99
Sports Reportage	0.10	1.83	0.19	0.19	2.89	1.25	1.40	2.21
Trad Narrative	1.85		1.17		6.93	0.10	5.08	2.05
Fiction	1.68		1.58	0.11	8.31	0.32	6.21	1.79
Formal Prose	5.51	0.10	3.05		8.07	0.39	7.28	1.08
News Scripts	3.27	0.30	4.36	1.09	8.92	1.98	11.69	3.57
Popular Writing	3.59	0.21	1.64	0.82	4.52	0.62	5.24	2.05
Average	2.20	0.72	1.65	0.53	5.90	0.67	5.28	1.95

Once again, the news script register seems to be the most nonconcordant in terms of the genitive marking of masculine nouns. However, it also has the highest level of concordance for all but masculine plural definite genitives. By making reference to the ratio of concordant to nonconcordant items per 1000 words, it should be possible to disambiguate these findings:

Table 98: Noun concordance data: feminine, singular, definite, genitive

Sports Reportage	2.67	Concordance ↓
Radio Interview	3.50	
Conversation	14.31	
News Scripts	14.44	
Trad Narrative	14.65	
Formal Prose	16.15	
Fiction	22.05	
Popular Writing	29.81	

Table 99: Noun concordance data: feminine, singular, indefinite, genitive

Radio Interview	3.00	Concordance ↓
Conversation	3.15	
News Scripts	3.69	
Trad Narrative	5.51	
Fiction	6.11	
Popular Writing	6.81	
Sports Reportage	7.21	
Formal Prose	21.33	

Table 100: Noun concordance data: masculine, singular, definite, genitive

Sports Reportage	2.31	Concordance ↓
News Scripts	4.51	
Popular Writing	7.29	
Conversation	10.68	
Radio Interview	11.86	
Formal Prose	20.69	
Fiction	25.97	
Trad Narrative	69.30	

Table 101: Noun concordance data: masculine, singular, indefinite, genitive

Sports Reportage	0.63	Concordance ↓
Radio Interview	1.71	
Conversation	2.18	
Trad Narrative	2.48	
Popular Writing	2.56	
News Scripts	3.27	
Fiction	3.47	
Formal Prose	6.74	

The trend that emerges from these tables is that sports reportage tends to be amongst the most nonconcordant registers – in terms of the genitive – followed by radio interview, while the more concordant (i.e. conservative) registers are formal prose and fiction, with traditional narrative the most concordant of the spoken registers. Averaging all of these scores together into one main table of genitival concordance (Table 102), we find the following:

Table 102: Average C: N ratio for genitive nouns

Sports Reportage	3.21	Concordance ↓
Radio Interview	5.02	
News Scripts	6.48	
Conversation	7.58	
Popular Writing	11.62	
Fiction	14.40	
Formal Prose	16.23	
Trad Narrative	22.99	

Interestingly, traditional narrative has leapt to the most concordant position. This is because of the outlier in the masculine singular definite table (Table 100) above, where it had a ratio of 69.30. A look at the actual lexical items in this subcorpus reveals that one sixth of the items are repeats. For instance, the phrase *Fear a' Churacain Ghlais*, a perfectly formed genitive, is repeated 11 times. However, the distribution of the other registers is as one would expect them overall.

7.1.4 Frequency of Nonconcordance per Tag Type

Although slightly digressive to the thrust of the analysis, if we examine, in the corpus as a whole, the nonconcordant data only in descending order of occurrence, we find the following:

Table 103: Nonconcordance per tag type

Tag	Average per 1000 words	Key
MSIG	1.95	• (M)asculine
FSIG	0.83	• (F)eminine
MPDG	0.72	• (S)ingular
MSDG	0.67	• (P)lural
FSDG	0.54	
MPIG	0.47	• (I)ndefinite
FSDD	0.36	• (D)efinite
FSID	0.34	
MSIN	0.24	• (N)ominative
MSDN	0.23	• (D)ative
FPIG	0.23	• (G)enitive
MSID	0.20	
MSDD	0.19	Tag Syntax:
FPDG	0.19	Gender–Number–Definiteness–Case
FSDN	0.17	

Confirming expectations harboured by many Gaelic users, indefinite singular nouns in the genitive have the highest frequency of nonconcordance (indefinite plural genitives are further down, but are still in the top half of the list). Second, are definite nouns in the genitive case (except for feminine plurals). Third, feminine nouns in the dative are grouped together, followed by, fourth, masculine singular nouns in the nominative. Finally, singular masculine nouns in the dative, plural feminine nouns in the genitive, and singular, definite, feminine nouns in the nominative are the least frequent of the nonconcordant forms found. Of course, as mentioned above, these levels are at least in part due to overall occurrence of the forms, regardless of whether they are grammatically concordant. However, the consistency is striking nevertheless.

7.1.5 Adjectival Case Marking

To investigate the levels of adjectival concordance in the corpus, first an overall measurement of concordant and nonconcordant adjectives was taken. A ratio was then calculated upon these scores. The results of this analysis are in Table 104:

Table 104: Adjectival grammatical concordance

Register	Ratio	Conc	Unconc
Sports Reportage	5.04	12.60	2.50
News Scripts	5.86	21.50	3.67
Radio Interview	8.01	12.90	1.61
Popular Writing	15.47	23.83	1.54
Fiction	19.66	26.93	1.37
Conversation	19.69	13.98	0.71
Trad Narrative	26.10	12.79	0.49
Formal Prose	38.49	26.56	0.69

Concordance



The rows have been sorted in ascending order based upon the ratio of concordance and nonconcordance as above. In many ways, Table 104 is similar to those dealing with the nominal data. Sports reportage is again the least concordant register, while formal prose and traditional narrative are amongst the most concordant registers. The news scripts are clearly amongst the most nonconcordant registers while fiction has taken a place in the middle, relative to where it was in the genitive noun data. The overall picture still describes a general correlation between concordance and production circumstances. It is possible that additional data would reallocate Table 104 slightly more in the direction of those above.

7.1.6 Summary of Noun Phrase Grammar

In summary, this section peered closely at noun phrase grammar, breaking down the results of §6.2.4.2 concerning the distribution of nonconcordance in the corpus. The findings indicated that most of the cases of nonconcordance were accounted for by the genitive case, in particular when the noun was indefinite. Formal prose and traditional narrative were consistently the most concordant registers while sports reportage, which has the most constrained production circumstance in the corpus, was the most nonconcordant. As a general rule, it appears that registers with more editing and fewer time constraints tend to be more grammatically concordant while those that are more spontaneous tend to be more nonconcordant. However, traditional narrative is a spoken, albeit rehearsed, register and has one of the lowest rates of nonconcordance while news scripts are written, and afford some time, if only minutes or seconds, for editing. The discrepancy between these two register types is interesting and, as is claimed in the conclusions (Chapter 9), probably due to the difference between their producers, along with effects related to translation in the case of news scripts.

7.2 Noun Phrase Complexity and Function

One of the major differences found in the statistical results was that features involved with incorporating information within the noun phrase were less frequent in spontaneous speech than other language varieties: complex nouns; prepositional phrases; attributive adjectives; genitive nouns; and – for conversation but not radio interview – relative clauses. Chafe (1982), Biber (1988), and Miller and Weinert (1998) have all obtained similar findings and it seems that these rest with both the less informational orientation of most spontaneous speech and the impact of unplanned production on the capacity for integration. Jucker (1992) found that the level of complexity in subject noun phrases was a good estimate of upmarketness in British newspapers. Thompson (1988) found that both English and Mandarin Chinese speakers in general tend to avoid heavy NPs in subject position. Miller and Weinert (1998: 139) pose the hypothesis that such a tendency is the result of processing constraints.

In this section, the feature ‘complex noun’ is broken down further. As mentioned in §6.2.3, for tagging purposes a complex noun was defined by the presence of more than one kind of nominal modification, between genitives, attributive adjectives, prepositional phrases, and relative clauses. In the coding for the current section, comparatives (e.g. *duine nas motha* ‘a bigger man’) and modifying numerals were classified with attributive adjectives while superlatives (e.g. *a’ chlach as truime* ‘the heaviest stone’) were grouped with relative clauses. Here, we will examine the frequency of the different constituents and consider the syntactic function of the complex NPs in the data set.

7.2.1 Data Coding

LinguaStat was used to retrieve all instances of the tag ‘complex noun’ for the two spontaneous spoken registers and for formal prose, which has been commonly used in the

literature to contrast informal speech. These hits were then separated into several different constituent categories. Miller and Weinert (1998), enlisting past research on the frequency of noun phrase constituents (Coulthard and Robinson 1968; Hawkins 1969), posit a ranked scale for dealing with their nominal data from English, Russian and German. This scale, from lowest rank to highest rank, runs as follows (pp. 144–145): 1) personal pronoun; 2) bare noun; 3) determiner + noun; 4) (determiner) adjective (adjective) noun; 5) (determiner) noun prepositional phrase; 6) X noun relative clause; 7) (determiner) noun participial phrase. In their work, these categories are mutually exclusive from one another; a noun phrase modified by both an adjective and a prepositional phrase would be counted only as having a prepositional phrase. Because the Gaelic data had already passed through one filter – having already been grouped as ‘complex’ – this approach needed to be tailored slightly. It was decided that the same general ranked scale would be used, but that it would be both attenuated and slightly extended. Because all of the data was already ‘complex’ the first three categories above (personal pronoun, bare noun, determiner + noun) were not applicable. Also, since there were always two or more modifiers for each noun phrase in the data set, one type of modifier would have to be chosen in preference to the other – that is, unless there were two of the same kind (e.g. two adjectives), where it seemed important to declare this. A further difference was that N-N genitives were included as one of the constituent categories in the current research.

The rank system decided upon resembled a game of poker with each deal having two or more ‘cards’ (i.e. features). If a complex noun tag showed a noun modified by both a prepositional phrase and an adjective, the prepositional phrase was counted only. However, as in poker, sets of the same features were given more weight; if there was a relative clause but also two adjectives, the adjectives were counted as important but not the relative clause. Although genitives were included, a quick scan of the frequencies indicated that they were more common than attributive adjectives, so they were given the lowest weight. Therefore, the ranked scale was ordered as follows, from lowest rank to highest: 1) one attributive adjective; 2) one prepositional phrase; 3) one relative clause; 4) two genitives; 5) two attributive adjectives; 5) two prepositional phrases; 6) two relative clauses. Necessarily, whenever one attributive adjective is counted, it is with a genitive as well, which has the lowest rank and cannot be counted on its own.

In addition to examining constituent frequencies, this data was used to investigate the syntactic functions of the complex noun phrases in the two language varieties. Six different positions were considered: subject of a transitive verb; subject of an intransitive verb; transitive object or subject of a passive/impersonal verb; object of preposition; genitive noun; and extraclausal. Noun phrases were considered extraclausal when they occurred in a detachment construction, in appositional phrases, or in parenthetical phrases.

7.2.2 *Tabular Data*

There were notable tendencies and differences in constituent structure for NPs in spontaneous spoken registers and formal prose, as seen Table 105 below:

Table 105: Constituent frequencies for complex NPs

<i>Constituent</i>	<i>Formal Prose</i>		<i>Spont. Speech</i>	
	<i>Count</i>	<i>Frequency</i>	<i>Count</i>	<i>Frequency</i>
1 Adj.	23	20.35	8	12.12
1 PP	29	25.66	23	34.85
1 Rel Clause	19	16.81	7	10.61
>1 Genitive	4	3.51	1	1.52
>1 Adj	29	25.66	24	36.36
>1 PP	7	6.19	1	1.52
>1 Rel Clause	2	1.77	2	3.03
Total	113	100	66	100

Most apparent in this Table 105 is the tendency in both categories for adjectives to co-occur rather than to be associated with another kind of modifier. This was stronger in spontaneous speech than in formal prose. There were no enormous differences between the two register types, but there were notable divergences: relative clauses, overall, made up a lower proportion of complex NPs in spontaneous speech than in formal prose; there was a slightly greater proportion of complex noun phrases with more than one PP in formal prose; and finally there were more cases of 1 adjective (plus 1 genitive) in formal prose but, a higher proportion of cases having >1 adjectives in spontaneous speech. The latter finding is notable as it diverges from the data of Miller and Weinert (1998). It is important to note, however, that this is not to say that there were more two adjective NPs in the spontaneous spoken registers, but only that they occurred more often than other types of complex nouns. Overall, although most of these statistics are not remarkable in and of themselves, when the examples drawn from the texts below are considered, the qualitative differences between the two conditions of texts will become very clear.

The data concerning syntactic function show more striking differences. In particular, as mentioned above, previous research has indicated that spontaneous speech tends to avoid heavy noun phrases in subject position, especially in transitive clauses. As seen in the following Table 106, there are instances evinced in formal prose, but none in spontaneous speech:

Table 106: Syntactic functions of complex NPs

Function	Formal Prose		Spont. Speech	
	Count	Frequency	Count	Frequency
Subject _{TRANS}	4	3.48	0	0.00
Subject _{INT}	20	17.39	11	18.03
Object _{TRANS}	13	11.30	7	11.48
Object _{PREP}	61	53.04	25	40.98
Genitive	6	5.28	1	1.64
Extraclausal	11	9.57	17	27.87
Total	115	100	61	100

Interestingly, there seems to be little difference between the varieties for intransitive subjects and transitive objects, but a difference did occur in the proportion of complex NPs in genitive positions and as objects of prepositions, with both more frequent in formal prose. The clearest difference, however, obtains in the proportion of extraclausal complex NPs. For spontaneous speech, this is the second most favoured position for complex NPs. This involves the tendency in speech for the separation of role and reference (Lambrecht 1994): heavy NPs are generated outside of the clause to which they pertain, while inside the clause they are referred to with a pronoun. This is seen in the examples to come and is most apparent as the number of NP modifiers increase in speech.

7.2.3 Examples from the Text

7.2.3.1 Single Adjective Complex NPs

Table 107 lists all of the instances of single adjective complex noun phrases in the spontaneous spoken sample. As we see, there is a very narrow range of vocabulary items: the word *m(h)òr(a)* 'big' occurs twice, *math* 'good' is common, and one phrase *ceann shìos an taighe* 'the far end of the house' is repeated three times. In addition, none of

these examples function as the grammatical subject. The written examples in Table 108 show a much wider word range and a greater level of noun phrase integration overall. Two of the adjectives, *dearg-fhuileach* 'red-blooded' and *soch-mhalairteach* 'marketable(?)' would very rarely, if ever, occur in spontaneous conversation. Moreover, some of the examples contain other nouns which are themselves modified. Example 3 has two complex NPs in the same clause: *astar math throighean sìos* 'a good distance of feet down' and *clachan garbha a' chladaich* 'the rough stones of the shore'. In example 5, *mòr-roinn* 'province', which is in genitival relation with the noun *roinn* 'part', is in turn the head of another genitive noun *Lochlainn* 'Norway'. Examples 14 presents an adjective-modified noun with a genitive modifier which itself takes an adjective: *bunaitean ùra a' chreidimh shoiseulaich* 'the new foundations of the Evangelical religion'. This is also the case in examples 16, *sheann ghnathas nam marbhrann gaisgil* 'the old customs of the heroic elegy', and 20, *deas làimh an airm Sheumasaich* 'the right hand of the Jacobite army'.

Table 107: Single adjective noun phrases from spontaneous speech

1. nise dè tha sinn a' feumachadh airson *sgioaba math* tug-of-war
'a good tug-of-war team'
2. ach feumaidh tu a h-uile seòrsa duine fhaighinn
'every kind of person'
3. thug mise luchd mhòr feòir gu [ainm]
'a big load of grass'
4. ach 's e taighean mòra nan daoine Sasannaich [sic] na Toffs
'the big houses of the English people'
5. ach bidh rùm aig ceann shìos an taighe
'far end of the house'
6. bidh sitting room aca aig ceann shìos an taighe
'far end of the house'
7. bidh iad a' suidhe aig ceann shìos an taighe
'far end of the house'
8. tha dà thaobh a' rathaid mhòir ann
'two sides of the big road'

Table 108: Single adjective noun phrases from formal prose

1. an aon seòrsa stòireannan agus sgeulachdan measail aig an t-sluagh air fad
'the same kind of stories and tales'
2. grunnan de thighearnan mòra fearainn sa cheann sin
'big lords of land'
3. a bha a' sìneadh gu cas astar math throighean sìos gu clachan garbha a' chladaich
'good distance of feet ... the rough stones of the shore'
4. gu math coltach ri sluagh eile na dùthcha
'other people of the country'
5. na Lappaich, à roinn a tuath mòr-roinn Lochlainn
'north part of the province'
6. Tha mac-an-duine ag àiteachadh iomadach seòrsa àite
'many types of place'
7. gus a shlighe a dhèanamh dhachaigh gu taobh siar Leòdhais
'west side of Lewis'
8. gu robh tricead nan seòrsachan fala sin eadar-dhealaichte anns na treubhan fà-leth
'those blood types'
9. Dòmhnall MacFhionghain, ceud Ollamh nan Cànan Ceilteach an Oilthigh Dhùn Èideann, air litreachas nan Gàidheal
'the first Professor of the Celtic Languages'

10. tha e a' dèanamh iomraidh air bàrdachd *chràbhach* na Gàidhlig
'religious poetry of Gaelic'
11. gun d'rinneadh sgrùdadh sam bith as fhiach ainmeachadh air bàrdachd *chràbhach*
na Gàidhlig
'religious poetry of Gaelic'
12. nach robh buaidh mhòr aig an t-seòrsa bàrdachd *seo*
'this type of poetry'
13. 's tha sinn a' faicinn an dàin mholaidh *seo* gu ìre mhòir a' call a lùths
'this praise poetry'
14. Agus seo an dearbh àm anns a bheil bunaitean *ùra* a' chreidimh shoisgeulaich
'new foundations of the religion'
15. mura robh meadhon brosnachaidh *eile* ann a thuilleadh air seo
'another medium of encouragement'
16. Bha Rob donn air ceum, barrachd is ceum, a thoirt air falbh bho *sheann* ghnathas
[sic] nam marbhrann gaisgeil
'the old customs of the heroic elegy'
17. Gabhaidh ceallan *dearg-fhuileach* mhic-an-duine an dealachadh bho chèile a rèir
nàdur nan stuathan a tha iad a' giùlain air an uachdair
'red blood cells of man'
18. mar a thig ailealan *ùra* am follais aig an lùcas ghineachan *seo*
'this locus of genes'
19. gus leasachadh a dhèanamh air frithfhiadh a' mhorair
'the deer forest of the lord'
20. agus bha iad air *deas* làimh an airm Sheumasaich
'right hand of the Jacobite army'
21. gur e taobh-an-iar na dùthcha - ceàrnaidhean *seann* Rìoghachd nan Eilean - as
fhaide leam
'the old Kingdom of the Islands'
22. nach tig latha *mòr* a' bhreitheanais orra fhathast
'the great day of judgement'

7.2.3.2 Single Prepositional Phrases

In Table 109, we have noun phrases modified by a single prepositional phrase in the spoken texts. Amongst the notable features here are the number of complex noun phrases occurring in extracausal positions. Examples 3 and 4 are in a detached construction and 5 and 16 occur as repetition-elaborations. Example 6 is interesting, exhibiting the kind of fragmentation that commonly occurs in spontaneous speech: *tòrr dhe na h-àiteachan mar sin iomallach* 'many of the places like that - remote'. Here, the prepositional phrase actually occurs between the head noun and its adjective modifier. It is as if the adjective *iomallach* 'remote' is an afterthought, but it is clearly meant to modify *tòrr* 'many'.

Many of the prepositional phrases in the spoken examples consist of only a single prepositional pronoun (7: 8, 13, 15, 21, 22) or a demonstrative (19, 20). Looking at the formal prose examples in Table 110, we see that none of the PPs are as simple as this; all are associated with full NPs. Furthermore, while slightly over one half contain NPs which are themselves modified (exx. 1, 3, 5, 6, 8, 9, 11, 15, 17: 19, 21, 23, 25, 28, 29), this is true for only one third of the spontaneous spoken examples (exx. 2, 10, 11, 14, 16, 17: 23). One impressive example from Table 110 has a prepositional phrase with two NP heads, the second of which is modified by two adjectives and a further prepositional phrase (ex. 17): *bàrdachd èibhinn mu dheidhinn trèanachan uilbheisteach agus baothalas coimheach, càl-shleamhnach mar òl tèa* 'humorous poetry about monstrous trains and strange, 'back-slipping' frivolities such as tea drinking'.

Another clear difference between the spoken and written examples again involves lexical range. The spontaneous spoken examples use only five simple prepositions (*aig* 'at', *de* 'of', *do* 'to', *le* 'with', and *mar* 'like') while the formal prose examples use eight simple prepositions (*à* 'out of', *aig* 'at', *air* 'on', *ann* 'in', *de* 'of', *eadar* 'between', and *fo* 'beneath') and two complex ones (*a thaobh* 'regarding', *mu dheidhinn* 'about'). One final observation is that there is a case of a heavy NP as a transitive subject (ex. 24) in the formal prose examples.

Table 109: Noun phrases modified by a prepositional phrase in spontaneous speech

1. chan e ach fear a ghabh am factaraidh mhòr [sic] *aig Kintorran*
'the big factory at Kintorran'
2. bhiodh Eachann e fhèin a' faighinn an ceathramh cuid *dhen tractor mhòr*
'the fourth part of the big tractor'
3. ach feumaidh tu a h-uile seòrsa duine fhaighinn na daoine mòra *aig an deireadh*
'the big people at the end'
4. ach feumaidh tu a h-uile seòrsa duine fhaighinn na daoine mòra *aig an deireadh na daoine beaga aig an aghaidh*
'the little people at the front'
5. chan eil ach dìreach bòrd mar sin square mòr *de bhòrd ann*
'a big square of table'
6. chan eil agad ach coimhead air Sealltain air an Nìribhidh PEI thall ann an Canada tòrr dhe na h-àiteachan *mar sin iomallach*
'many of the places like that remote'
7. ach 's e taighean mòra nan daoine Sasannaich [sic] na Toffs na Toffs tha sin-ach na taighean mòra *aca*
'their big houses' [Lit. 'the big houses at them']
8. cò an coach aig an aon taobh *agaibhse*
'your one side' [Lit. 'the one side at you (emphatic)']
9. a bheil ciall sam bith a's a' leithid seo a *bheachd* a Sheonag
'this type of opinion'
10. bha ceithir acraichean fichead againn *de thalamh dubh*
'we [had] twenty four acres of black land'
11. air bàrr nan cruachan no rudeigin luideach *dhen t-seòrsa sin*
'something silly of that sort'
12. gum feumar sùil a thoirt air seòrsa eile *de dhaoine*
'another type of men'
13. 's tha sinn ga iarraidh son na Pàrlamaid Nàiseanta againn *fhìn*
'our own the National Parliament' [Lit. 'the National Parliament at us ourselves']
14. se crìochan gu math sònraichte a bha sa h-uile gin *dhe na ceathramhan sin*
'every one of those quarters'
15. chan eil ann ach pìos beag *dhe* gun teagamh
'a small piece of it'
16. cuine a bha a leithid seo a' tachairt leithid seo *de lathaichean* [sic] *ro Bhealltainn no an dèidh Bhealltainn*
'these kind of days before Beltane or after Beltane'
17. agus tha iad a' dol an lùib agus mar gum biodh feilltean no crìochan sònraichte *dhen bhliadhna againn*
'festivals or particular borders of our year' [Lit. '... of the year at us']

18. sin is coireach ris an eadar-dhealachadh eadar a' bhliadhna ùr *mar a tha i againn an-diugh* agus an t-seana-bhliadhn' ùr
'the New Year as we have it today'
19. tha mi fhìn an aghaidh rud sam bith *mar seo* a' tachairt
'anything at all like this'
20. sam faod sinn a bhith 'leagail le armachd cumhachdach *mar seo* air terrorists
'powerful armaments like this'
21. gu bheil aobhairean [sic] aig na h-Arabaich a bhith ag iarraidh àite tàibh *dhaibh fhèin*
'a living place of their own'
22. ò fhuair sinn rèidh air deannan math *dhiubh*
'a good group of them'
23. an e sin an seann bhoireannach *leis am falt* [sic] *gheal*
'the old woman with the fair hair'
24. gun robh e am fear sin *dhiubh* an t-athair do bhean Thormaid an Gobha na MacDonalds sin
'that one of them'

Table 110: Noun phrases modified by a prepositional phrase in formal prose

1. fhad agus a mhair an ceangal dlùth *eadar an dà dhùthaich*
'the close tie between the two countries'
2. Agus seo an dearbh àm anns a bheil bunaitean ùra a' chreidimh shoisgeulaich gan suidheachadh air Ghàidhealtachd agus seòrsa ùr *de mharbhrann* a' nochdadh
'this new type of elegy'
3. Cha do chuir Iain MacCoinnich ach seachd rannan *de 'Ghaoir nam Ban Muileach'* ann an Sàr Obair
'seven verses of "The Cry of the Mull Women"'
4. gus an tugadh bho mhuinntir a' Bhràighe cead feuraich *air Beinn Lì*
'permission of grazing on Ben Li'
5. Ach gabhaidh na sluaghan a bhuineas do'n dà threubh sin dealachadh a thaobh tricead iomadh gine *anns an dà shluagh*
'frequency of many genes in the two populations'
6. Tha buaidh aig a' choimhearsnachd air na sluaghan fa-leth *de gach seòrsa*
'the separate populations of each type'
7. an aghaidh achdan an t-siamarlain an 1874 a cheud aon dhe na còmhstrithean sin a *thaobh fearainn*
'those uprisings pertaining to land'
8. tha bàrdachd Ghàidhlig *de dhà bhun-sheòrsa* 'ga dèanamh an diugh
'Gaelic poetry of two fundamental types'
9. a tha ann an aon dual *de'n nua-bhàrdachd*
'in one dimension of new poetry'
10. rinneadh tomhas mòr *de dhearmad* air a' bhàrdachd spioradail
'a great measure of omission'
11. tha romham sa phàipear seo sùil nas dlùithe a thoirt air aon seòrsa *de dhàn spioradail*
'one type of spiritual verse'
12. bho rinn an t-Oll. Iain MacAonghais mìneachadh lèirsinneach air gnàth-bhriathrachas, no reatoraic, an dàin mholaidh *sa Ghàidhlig*
'speech-mode, or rhetoric, of the eulogy in Gaelic'
13. agus b' e sin na marbhrannan moralta aig *Rob Donn*
'the moral elegies of Rob Donn'
14. cha robh àite anns an leithid sin *de shuidheachadh* do dhreuchd an lèirmheasaiche
'that type of situation'

15. gu h-àraidh ann am bàrdachd nan eilthreach anns na bailtean mòra
'the poetry of the emigrants in the big towns'
16. ann an cus earbsa à comhriaghail
'in too much reliance of formulae'
17. dòighean mar mhòr-amharas air nithean ùra (bàrdachd èibhinn mu dheidhinn
trèanachan uilbheisteach agus baothalas coimheach, cùl-shleamhnach mar
òl tèa
'humorous poetry about monstrous trains and strange, 'back-slipping' frivolities
such as tea drinking'
18. tha an t-eilean roinnte eadar a' chuid sin de a shaoghal
'that type of its world'
19. Chaidh na bun-chomharraidhean aig an riaghailt sin de sheòrsachan fala ris an
can sin
'that rule of blood types'
20. Tha ciudad àrd de na Sìnich aig a bheil
'a high percentage of the Chinese'
21. Coltach ri sluagh Innis Tìle, agus sluaghan eile air crìochan na Roinn-Eòrpa
'other populations on the border of Europe'
22. ach air an latha ud, an naodhamh latha deug de 'n Ghiblinn
'the nineteenth day of April'
23. a fhuaras air iomall duilleig ann an leabhar air choreigin
'border of a page in one book or another'
24. Ach tha leithid seo de thuaraisgeul a' mealladh dhaoine
'this type of description'
25. gu bheil feum air sùil eile a thoirt air ainmeannan-àite an ceann an deas na
rìoghachd
'place names in the southern reach of the kingdom'
26. tha e dualach dhan a h-uile cànan fon ghrèin
'every language under the sun'
27. a' cinneachadh fada roimhe sin an corra chèarna de Alba
'the odd district of Scotland'
28. anns a' cheud leth de'n ochdamh linn deug
'the first half of the eighteenth century'
29. dh'eug am fear mu dheireadh de na seanchaidhean foghlumaichte seo
'the last one of these educated tradition bearers'

7.2.3.3 Single Relative Clauses

Turning to the relative clauses, one again finds a situation of greater complexity overall in the formal prose. To begin with, most of the items in the spoken examples in Table 111 may even be questionable inclusions. There is an established tendency in ScG for demonstratives to be associated with tense and, in most of the examples – 1, 4, 5, and 6 – we find this kind of construction. Example 7 can also be included; it uses *ann* in a demonstrative sense, meaning *there* or *here*. Semantically, there is little difference between saying *air na diofar sgiobaidhean seo* 'on these different teams' and inserting a tense marker as in example 1. In part, this usage may serve to break up the utterance in the way so characteristic of speech, to ease processing on the part of the recipient. If we examine the examples in Table 112, we see that none of them evince this kind of construction.

Example three in Table 111 approaches the kind of complexity observed in written language, with a long relative clause containing three verbal nouns and an object. This example comes from a panel discussion on Middle East politics and it is possible that the greater formality of the context makes such examples more likely. Looking at Table 112, however, one sees a greater overall complexity than in the spontaneous spoken examples. First of all, one finds two preposition-headed ('pied-piped') relative clauses in examples 7

and 15. Also, there is a superlative in example 3 and a copular relative in example 11. Example 10 is particularly complex, even for writing: here, the subject of the relative clause (*Dòmhnall MacFhionghain* 'Donald MacKinnon') is further modified by a long appositive, the head noun of which is itself modified by an adjective, a genitive, and a prepositional phrase. Such examples are facile enough in writing but would be intolerably heavy for the rapid-fire pace of most spontaneous speech.

Table 111: NPs modified by a relative clause in spontaneous speech

1. nuair a bha sinn a' bruidhinn air na diofar sgiobaidhean *a tha seo* an-dràsta
'these different teams that are here'
2. sin an aon airgead *a bha tighinn a-staigh dhan an dachaigh* fad na bliadhna
'the only money that was coming into the household'
3. agus tha Reagan a-nis a' feuchainn ri cron a dhèanamh air na daoine sin *a tha cuideachadh agus ag ullachadh agus a' teagasg na terrorists*
'those people who are helping and readying and teaching the terrorists'
4. goirid air a bhith 'losgadh orra 'san dòigh eagalach *a th' ann an seo*
'in the frightful way that is here'
5. nach biodh tu cho math toll a dhèanamh leis an digger is fear dhe na windmills
bheag tha siud a chuir air
'the little windmills that are there'
6. nan slaodadh e fear dhe na trailing ploughs mhòr *a tha siud*
'the big trailing ploughs that are there'
7. nach eil càil a cho-pàirt ac' a's an obair uabhasach *a th' ann*
'the horrible work that is there'

Table 112: NPs modified by a relative clause in formal prose

1. sgeulaiche barraichte *a bha a' fuireach* anns an Àird Mhòir
'an exceptional storyteller who was living'
2. an uair a bha cuid de na h-uaislean Èireannach *a tha air an ainmeachadh* innte
beò fhathast
'the Irish gentry who were named'
3. 's e seo nam bheachd-sa an leabhar *as luachmhoire* mu litreachas na Gàidhlig
'the most valuable book about Gaelic literature'
4. agus cuideachd mu dhol-sìos a' choluadair ghaisgeil *a bha mar bhonnstèidh*
'the heroic company that was as a foundation'
5. air cùl nan atharrachaidhean soch-mhalairteach *a bha a' sìor sgapadh*
'the Socio-economic changes that were continually spreading'
6. gu bheil cruth air a' Ghàidhlig an Alba *a bheil cruth na Cuimrig* [sic] an
cuimhne duine sam bith
'shape on the Gaelic in Scotland that brings the shape of Cumbric'
7. an cuimhne duine sam bith *aig a bheil eòlas* air na dhà
'any person at all who has acquaintance'
8. nach d'fhàg an luchd-leanmhainn cumanta Gàidhealach lorg domhainn no làrach *a tha furasda aithneachadh*
'a deep trace or remain that is easy to recognise'
9. dh'iarradh air a' bhalach òg *a bha ri faire*
'the young boy that was at watch'
10. anns na h-aisteachan liomhte *a sgrìobh Dòmhnall MacFhionghain*, ceud
Ollamh nan Cànan Ceilteach an Oilthigh Dhùn Eideann, air litreachas nan
Gàidheal
'the polished essays that Donald MacKinnon wrote'

11. gun d'rinneadh sgrùdadh sam bith *as fhiach ainmeachadh*
'any research at all worth mentioning'
12. bha an tomhais soirbheachais a bha gu bhith le bàrdachd an crochadh, anns a' cheud àite, air
'the measure of success that was to be with poetry'
13. gu bheil barrachd sliochd ga fhàgail leis an fheadhainn sin a tha comasach
'those ones who were able'
14. chaidh mòran de rannsachadh a dhèanamh air tricead gineachan nan seòrsachan fala a tha aig mac-an-duin
'the blood types that belong to mankind'
15. Chaidh na bun-chomharraidhean aig an riaghailt sin de sheòrsachan fala ris an can sinn
'blood types which we call'
16. Ach ma choimeasar na seòrsachan fala ABO a tha acasan
'the blood types ABO that they have'
17. Cha do ghabhadh mòran suime de an cuid adbhaircan na idir de gach casaid a rinn iad
'every case that they made'
18. tha an cànan cho beairteach fhathast is gu bheil iomadh cùil is cèal dhi a tha ri an rannsachadh
'many a nook and secret of it that has yet to be researched'
19. gu bheil na faclan a' nochdadh dualchainnt shonraichte a tha Albannach seach Èireannach
'a particular dialect that is Scottish as opposed to Irish'

7.2.3.4 Two or More Genitives

There is only one example of a double genitive in speech: *le cloinn Theàrlaich Dhòmhnail a' chidhe shìos*. Personal names and place names are often as complex as this, but they are fossilised to a large extent and cannot be considered productive instances of noun phrase construction. In contrast, the written examples below show a similar level of complexity, but all evince conjunction of the other genitive nouns. For instance, in example 1, *seòrsachan* is the head noun followed by the genitive form of *fuil* 'blood' – *fala* – which itself is modified by *ABO*, a blood type. The phrase *seòrsachan fala ABO* forms a constituent which is then modified by no less than three genitives: *nan Innseanach*; *nan Hungairianach*; and *na Gypsidhean Hungairianach*. Interestingly, the first two genitives show the plural, non-labial definite article *nan* while the last reverts to the nominative/dative form *na*. There seems to be a tendency in the language for more distal genitives not to be marked as such.

Table 113: Multiple genitives in formal prose

1. co-dhiù, ma rannsaichear seòrsachan fala ABO nan Innseanach, nan Hungairianach agus na Gypsidhean Hungairianach
'blood types ABO of the Indians, the Hungarians and the Hungarian Gypsies'
2. ann an dòighean àraidh leithid cruth nan sùilean, no dath fuil is dath shùilean
'such as [takes genitive] the shape of eyes, or the colour of hair and colour of eyes'
3. thàinig sgrios cuideachd air saoghal nam filidhean agus nan seanchaidhean foghlumaichte a bha a' cosnadh am beò-shlàinte
'the world of the educated class and the educated tradition bearers'
4. ann am beul-aithris Ghàidheal Alba agus Eirinn
'the oral tradition of the Gaels of Scotland and of Ireland'

7.2.3.5 Two or More Adjectives

As mentioned above, there is a tendency for adjectives to co-occur rather than to be associated with other types of modifiers. This explains the larger number of returns in the tables below, showing more >1 adjective complex NP, than in those dealing with only one adjective. Some of the spoken examples are remarkable for their level of modification. The first example in Table 114, for instance, has a head noun modified by four adjectives: *mòra*, *tapaidh*, *làidir* and *fiadhaich*. Example 2, in addition to two adjectives, has a relative clause and a prepositional phrase. Example 17 also has two adjectives and a prepositional phrase. However, in these three examples, as well as examples 5, 6, 7: 8, 9, 14, and 22, the noun phrase occurs extracausally. These ten cases amount to 42% of the spontaneous spoken examples altogether. Looking at formal prose, in Table 115 only examples 3, 13, 29, and 30 are extracausal, amounting to only 13 of the occurrences. In addition, the formal prose cases are more complex overall: examples 10, 19, 23, and 26 evince NPs modified by two adjectives in addition to a relative clause while example 30 has two adjectives and a prepositional phrase. Furthermore, the complex noun phrase in example 6 is itself a genitive modifier of another noun phrase. These instances of integration are in contrast to the fragmentation found in the spoken examples. The extracausal nature of many of the occurrences has already been mentioned. The head noun in example 22, Table 114 looks as if it is the subject of a *'s* cleft with an elided copular verb. This is followed by the phrase, *beagan chaoirich*, which, while not syntactically linked to any of the preceding constituents, is a further comment on the noun *croitean*. Finally, the adjectives used in the spontaneous speech examples are, for the most part, very commonplace; *mòr* alone occurs twelve times. The only unusual occurrences are the comparatives *nas feallsdanaich* and *nas fhoghlaimaichte*, which occurred in the radio discussion on Middle East politics mentioned earlier. The adjectives in the formal examples show a greater variety and specificity. Some, such as *clasaigeach*, *eachdrachail*, *sòichealach*, and *cùl-sheamhnach*, not to mention *foghlaimaichte*, would be very uncommon in everyday parlance.

Table 114: NPs modified by two or more adjectives in spontaneous speech

1. an sgioba a b' fhearr a ghlaodh an sgìre ri chèile *daoine mòra tapaidh làidir agus fiadhaich* cuideachd
'big capable strong and fierce men'
2. 's ann aig a' choinneamh *sin eile* a bha ann an Loch nam Madadh air poverty air an oidhche sin
'that other meeting'
3. tha *dà bhàta mhòr* aca fhèin
'two big boats'
4. ach dh'fheumadh tu tràilear *mòr uamhasach* air a shon
'a big horrible trailer'
5. liobagan *mòra mòra* an fheadhainn ud
'big big flounder'
6. 's ann car cruinn a bha i spotan *mòra dearga*
'big red spots'
7. an rathad *mòr cianail seo* a tha iad a' togail an Sgalpaigh an-dràsta tha coinneamh agam a-màireach mu dheidhinn
'this big awful road'
8. factaraidh *mhòr mhòr chianail* a' dol suas
'a big big awful factory'
9. an e daoine *mòra brùideil* no dè
'big brutish men'
10. gun tèid sin air adhart airson *ceud bliadhna eile*
'another one hundred years'

11. bidh duilgheasdan *teicneageach agus eile* ann
'technical and other difficulties'
12. a fhuair mi ann an àite *sam bith eile*
'any other place'
13. far a robh daoine *òg is sean* is dhen a h-uile dath ann
'young and old people'
14. o chionn deich bliadhna 's còrr agus mi ann an Ìle an *seann* duine *sin* ag innse mu mhanadh
'that old man'
15. bidh fealla-dhà agus horo-gheallaidh eile againn ann am bliadhna *ùr eile*
'another new year'
16. 's e *fìor shean* rud a bh' ann
'a truly (adj.) old thing'
17. sàilleadh an t-*seann* bhliadhna *ùr* a's an t-seann chunntais cuin' a bha e a'
tòiseachadh
'the old New Year'
18. chan eil an dà bhliadhna *sin* a' tighinn a rèir a-chèile
'those two years'
19. agus feumaidh sinn saoilidh mise dòighean tòrr *nas feallsdanaich agus nas fhoghlaimichte* obrachadh a-mach
'ways much more philosophical and educated'
20. a bheil an cnatan air duine *sam bith eile* thall an-sin
'any other person'
21. suidh thall treiseag *bheag eile*
'another wee bit'
22. dè tha duine sam bith a' dèanamh orra croitean *beaga luideach* a tha thall an-sin
co-dhiubh beagan chaoraich
'small ragged crofts'
23. thàinig an t-uisge *mòr sin-ach*
'that big rain'
24. 's tha balla *mòr leathann* ann
'a big wide wall'

Table 115: NPs modified by two or more adjectives in formal prose

1. 'S ann air bàrdachd de'n *dàrna* seòrsa *sin*
'of that second kind'
2. on sgeulaiche *ainmeil nach maireann*⁴¹ Donnchadh Dòmhnallach
'the famous storyteller not living'
3. --agus cuid dhe na sgeulachdan *mòra eadar-nàiseanta*
'the big international stories'
4. dh'eug am fear mu dheireadh de na seanchaidhean *foghlumaichte seo*
'these educated tradition bearers'
5. chaidh na bha air fhàgail de na *seann* sgoileirean *seo* as an t-sealladh
'these old scholars'
6. a bha a' cosnadh am beò-shlaint ann an cùirtean nan *uaislean mòra Gàidhealach*
anns an dà dhùthaich
'big Gaelic gentry'
7. An uair a chaidh an *seann* saoghal *Gàidhealach* am mutha
'old Gaelic world'

⁴¹ The phrase *nach maireann* 'not living' is properly a negative relative clause with an obsolete simple present verb form, but it has been so crystallised in Gaelic that it seems appropriate to consider it adjectival.

8. a thaobh nan seachaidhean agus nam filidhean ionnsaichte a bha, nan cuid sgriobhaidh, a' cleachdadh an aon seòrsa de Ghàidhlig *chlasaigeach* agus *sheann-fhasanta*
'classical and old-fashioned Gaelic'
9. gu bheil an tràchdas seo mar dhòrnaig *bhig eile* air a' chàrn
'another wee handfull'
10. bha seo mar thoradh air tachartasan *eachdrachail* agus *politigeach* nach ruig mi a leas a rannsachadh an dràsda
'historical and political occurrences'
11. a bha ri a fhaighinn ann an iomadach cultur *Eòrpach* agus *Asianach* ann na Lintean Meadhonach
'many European and Asian cultures'
12. iomradh a thoirt air an sgoileir *ainmeil sin* Coinneach Jackson
'that famous scholar'
13. tha Oilthigh Dhun-Èideann, cleas *iomadach* acadama *eile* 's an latha-an-diugh
'many other universities'
14. nach d'fhàg an luchd-leannmhainn *cumanta Gàidhealach* lorg domhainn no làrach a tha furasda aithneachadh
'the common Gaelic contingent'
15. agus tha dà òran ann air an ainmeachadh air Mairearad: 'Gaoir nam Ban Muileach', agus duan *beag eile*
'another small poem'
16. bha an slaugh cumanta an sin dhen t-*seann* stoc *Bhreatannach*
'the old British stock'
17. gu bheil a' Ghàidhlig na cànan *foghlaimte, sgrìobhte* fad an rèis
'an educated, written language'
18. tha dà shluagh *eile* anns an Eòrpa an iar nach eil cinnt aig daoine co às a thàinig iad
'two other populations'
19. a' tighinn bho'n t-*seann* chànan *Innseanach*
'the old Indian language'
20. is iad fo smachd dà aileal *eadar-dhealaichte*
'two different alleles'
21. thàinig grunn de threubhan *daonna eadar-dhealaichte* am bith
'different human tribes'
22. a thaobh fearainn a dh'fhàs lionmhar anns na h-eileanan feadh an *ath dhusan* bliadhna
'the next twelve years'
23. an àite anns nach soirbhicheadh le caoraich no le ainmhidh *bhuannachdail sam bith eile*
'any other profitable animals'
24. agus mar sin 's e ealantas rannaigheachd agus geòire cainnte air an riochdachadh anns an fhacal *ghleusda, dhrùidhteach* air an robh an t-tinneas
'the crafty, impressive word'
25. Bha dlùth-cheangal eadar an traidisean agus comhdheilbh *sòichealach àraidh*
'a particular happy configuration'
26. mar a thuit an comhdheilbh *sòichealach seo* as a chèile
'this happy configuration'
27. 's far an robh an t-side gu math (laoich agus banaraich air fonn *aodharach, grianach*)
'a happy, sunny mood'
28. dòighean mar mhòr-amharas air nithean ùra (bàrdachd ìbhinn mu dheidhinn trèanachan uilbheisteach agus baothermal *coimheach, cùl-shleamhnach* mar òl tèa
'strange, 'back-slipping' frivolities'
29. Ach gabhaidh na sluaghan a bhuineas do'n dà threubh *sin* dealchadh a thaobh tricead iomadh gine anns an dà shluagh
'those two tribes'

7.2.3.6 Two or More Prepositional Phrases

There was only one example of an NP modified by more than one prepositional phrase in the spontaneous speech texts: *a' chlann nighean 's na balaich leis le eich air feadh na h-eilean bidh iad a' coinneachadh shìos ann an...* Once again, this is extracausal, a case of left-detachment. One of the formal prose examples in Table 116 is also extracausal (example but the rest are well integrated). As before, there is a variety of different prepositions represented, particularly for such a small number of examples: *de*, *fo*, *os cionn*, *ann*, *le* and *a*. Finally, many of the NPs within the prepositional phrases are themselves modified: example 1 has an adjective – *de litreachas sgriobhta*; example 3 has an NP modified by a genitive with adjective – *fo bhonn bruthaich àird*; example 4 has a genitive with adjective – *de ghnathanan [sic] nan seann marbhrann* – and a noun modified by an adjective – *de ghnathanan ùra*; finally example 6 has an adjective modifying *craiceann* – *dorch* – and a genitive modifying *muinntir* – *Cheann a Tuath Afraca*.

Table 116: NPs modified by two or more prepositional phrases in formal prose

1. *bha dà shruth toirteil – de litreachais [sic] sgriobhta agus de bheul aithris a' sior bheathachadh ... a chèile*
'two fertile strands – of written literature and of oral tradition'
2. *cha ruig sinn a leas a dhol nas fhaide na an treubh ainmeil sin de bhàird agus de sheanchaidhean, Clann 'ic Mhuirich*
'that famous clan of bards and of tradition bearers'
3. *cha robh ann ach ceum glè mhi-shocair fo bhonn bruthaich àird agus os cionn bearraidh*
'an unsteady step beneath a high brae and above a steep decline'
4. *a tha a' comharrachadh nam marbhrannan soisgeulach gu lèir, measgachadh de ghnathanan nan seann marbhrann 's de ghnathanan ùra*
'a mix of the ways of the old elegies and of the new ways'
5. *'S ann aca a tha an tricead as lugha anns an Roinn Eòrpa de sheòrsa fala B*
'the lowest frequency in Europe of the B blood type'
6. *bhiodh e duilich Arabach le craiceann dorch, a mhuinntir Cheann-a-tuath Afraca*
'Arab with dark skin, of the people of North Africa'
7. *'Se cur an aghaidh a' chrionaidh sa ann an cultur agus ealain*
'this withering in culture and arts'

7.2.3.7 Two or More Relative Clauses

Only two instances each of NPs modified by more than one relative clause were found in the spoken and written sample. In example 1 of Table 117 and both examples of Table 118, there is a superlative (e.g. *a b' fheàrr*) followed by a relative clause (e.g. *a ghlaodh an sgìre ri chèile*). Example 2 in spontaneous speech shows a demonstrative with tense marking (*a tha siud*) followed by a fuller relative clause (*a chosg fortan*). Obviously, this is a rare feature and a larger corpus would be needed to show a greater range of diversity and a clearer indication of what to expect in the two language varieties.

Table 117: NPs modified by two or more relative clauses in spontaneous speech

1. *an sgioba a b' fheàrr a ghlaodh an sgìre ri chèile*
'the best team that the district called'
2. *'s ann aca a tha an t-àite mòr ùr a tha siud a chosg fortan*
'the big new place that is there that cost a fortune'

Table 118: NPs modified by two or more relative clauses in formal prose

1. air cuid dhen Ghàidhlig *sine a lorgadh riamh ann an Alba*
'the oldest Gaelic that was ever found in Scotland'
2. gheibhear i, leis an tricead *as motha a th' aice air an talamh*, am measg nan Lappach
'the biggest frequency that it has on earth'

7.2.4 Summary of NP Complexity and Function

This section surveyed the noun phrase complexity of the formal prose and spontaneous speech (i.e. conversation and radio interview) subcorpora. Although there were few major differences in frequency for each type of complexity, the functions of the complex NPs were seen to contrast significantly. Most notably, spontaneous speech showed a marked tendency to have complex noun phrases in an extraclausal position. Also, there were fewer cases in spontaneous speech of complex NPs in the subject position of a transitive clause. The examples from the texts highlighted the qualitative differences between the two register types, with formal prose showing a greater lexical range, more 'upmarket' vocabulary items (e.g. scholarly adjectives), more examples of NPs in modifying constituents which are themselves modified, and a greater variety of construction types. The spontaneous spoken registers, in general, had simpler, more fragmented examples than the formal prose, although there were exceptional cases where modified NPs reached the complexity seen in formal prose.

8 Profiles of the Individual Registers

This section reviews the findings of chapter 5–7 in terms of the eight individual registers, each of which presents a different weighting on the linguistic features surveyed. By delineating the most characteristic of these features for each register and relating them to one another it is possible to begin to assess the conventional associations between context and linguistic form. This occurs through an interpretation of bundles of features correlating with one another, which are understood as sharing similar functions and indicative of certain contextual variables. Although the current research does not formalise the concept ‘feature bundles’ – or dimensions – in the way that Biber does through factor analysis, it is possible to inductively generate similar interpretations, which could be followed up empirically in future research.

After each feature mentioned in the text below, reference is made to where the register in question ranked upon it, given as a number between 1 and 8. A ‘{1}’ indicates that the register had the greatest return for that feature in the corpus, while ‘{8}’ indicates that it had the lowest. If the register’s score was tied with others in the corpus, this is indicated by a range, e.g. {1–2}. However, this is only a rough guide to distribution and it is important to consult the actual means and pairwise comparisons for concise interpretation. In general, reference will only be made to features that stand out as robust register markers, yet other features may be discussed from time to time for comparative purposes.

8.1 Conversation

Information structure and discourse-level features. Conversation displayed a high level of both *’s e* clefts {2} and demonstrative clefts {1}. As mentioned in §5.1, perhaps a good way of capturing the quality of a register like conversation, relevant to these kinds of features, is to say that it has more discourse freedom than other registers – more points of reference in the discourse or temporal-spatial context available for further comment. *’S e* clefts are often used in conversation for contrastive effect, such as in the second cleft (underlined) of example (105):

- (105) <a> *’s e leabhar gu math doirbh a tha sin a [ainm] tha mise a’ smaointinn gu bheil e doirbh fhios agad an leabhar aig [ainm] an oidhche roimhe*
 <c> *’s e leabhar glè mhath a th’ aig [ainm] an-dràsta ...*
 <a> it’s a very difficult book that that is [name-VOC] I am thinking that it is difficult you know the book [name] had the other night
 <c> it’s a very good book that [name] has right now ...

Here the first, identificational cleft reaches back to previous discourse to comment on a particular referent, a book. The second cleft introduces a different, but related referent and states a contrast with the previous referent – that it is a good book, not a difficult one. If an unmarked construction was used – *tha leabhar glè mhath aig [ainm] an-dràsta* – there would not be the same contrastive effect; it might be interpreted as more of a ‘beside’.

The most common discourse function of demonstrative clefts in conversation seems to be in ‘wrapping up’ a piece of discourse, such as in summarising a previous strain of discourse before moving forward, agreeing with something that a person has said, or stating that no more needs to be said on the matter. These qualities are reflected in some of the phrases that are common in this sub-corpus: *sin an rud* ‘that’s the thing’; *’s e sin a th’ ann* ‘that’s what it is’; *sin mar a tha e* ‘that’s the way it is’; *shin agad è* ‘there you have it’.

Interestingly, while *sin-* or *siud-*headed demonstrative clefts nearly always refer to previous discourse in one way or another, *seo-*headed demonstrative clefts, which are much rarer, seem to introduce a new topic or point forward in some way. This quality is seen in example (106), which is only one of two instances in the conversation texts *in toto*:

(106)

[4] *seo an t-àm a bhaite a' cuir dhan Eilean Caol* [sic] *airson nan liobagan*

[2] *àidh*

[3] *uill tha fear an-seo a bha mi a' feuchainn ri bhith 'smaoineachadh air 's tha t-èile ga chuir a-mach às mo chlaigeann*

[5] *chan eil ach tha mi ag ràdh 's fhad o chuala mi mu dheidhinn cha do thachair dad fhathast*

[2] *chan fhaigh thu iad co-dhiubh*

[4] *chan eil e ach nonsense tha mise 'dol dha fheuchainn co-dhiubh*

[5] *feuch e ma-thà*

[2] *dh'fheuch mise e uaireigin ma-thà nuair a bha [ainm] ris mi fhèin dh'fheuch mi shìos man Bhaile Shear an-shin is cha d'fhuair mi dad*

[4] *bhithinn-sa ga fhaighinn thall air chùl an Eilean* [sic] *Dubh* [sic] *ann an shiud-ach gu leòr dhiubh uaireigin*

[4] this is the time that they would set to the Slender Island for the flounder

[2] aye

[3] well there's this man that I'm trying to think of and another one is putting him out of my skull

[5] no but I am saying it's a long time I've heard about it and nothing has happened yet

[2] you won't find them anyway

[4] it's only nonsense I'm going to try it anyway

[5] try it then

[2] I tried once then when [name] was doing it me myself I tried down around Baleshare there and I didn't get anything

[4] I would be finding it over at the back of the Black Island over there lots of them once

Speaker 4 uses a *seo*-headed demonstrative cleft to introduce a topic change – fishing for flounder. The conversation has been previously about local fish merchants and the varieties of fish that they have on offer. After this topic shift, even after [3] interrupts, the conversation stays on the new topic for 18 turns.

Conversation, like most of the other registers, exhibits a overwhelming preference for postposing (.63) rather than preposing (.32) its subordinate clauses. Moreover, it had few returns for fronted adverbials and peripheral elements {7} which, like clefts, were characterised as bridges to previous discourse. However, there must be a distinction between this kind of emphasis and that found in 's *e* and demonstrative clefts. A good indication of this is that fronted adverbials and peripherals occur with the greatest levels in written texts.

Conversation had the greatest frequency of both left {1} and right {1} detachment. These features have been interpreted as indicating fragmentation, and serve to spread bits of information out over several clauses to aid in the comprehension of listeners.

Clause types. The clausal features showed that conversation tends to have few coordinated clauses {6} although this may increase with fewer participants, in proportion to the amount of narrative in evidence. In addition, clauses in conversation tend to be smaller, as indicated by the greater frequency of main clauses {2} in the register. There are few relative clauses {7}, as found in other 'on-line' registers such as sports reporting. Also, there were few complement clauses {7}, perhaps an indication of a lower informational component. Finally, there were relatively few concessive clauses {6}, which seems to be true of spoken registers in general.

Verbal morphosyntax. Conversation evinced the lowest occurrence of perfect aspect {8} in the registers. The reasons for this are unclear, but perfect aspect tends to be more prevalent in written registers. Copular verbs {1} were very common in conversation although the reasons for this are also unclear. Clausal negation {1} obtained most frequently in conversation, perhaps an indication of greater involvement in the register, manifested in more contention. It contrasts on this feature to registers with a more 'factual' orientation such as sports reportage and news scripts. Interrogatives {1} were also found with the greatest frequency in conversation, clearly a sign of its high interactivity. On the other hand, conversation had only the third highest level of imperative verbs {3}, behind traditional narrative and fiction. This may have had to do with the production domain; if the recordings had all been made in the kitchen, for example, whilst people were cooking food, perhaps a greater level would have been detected.

Tense on its own was not overly marked in conversation. The level of past tense {3} is a reflection of the amount of narrative found; high for a 'non-narrative' register. Present tense {2} seems to co-occur with high frequency in more involved registers and those which emphasise current reference.

In terms of valence reducing constructions, conversation evinced some of the lowest means for all three features: valence decreasing with agent present {6-8}; synthetic impersonals/passives {7}; and total number of passive clauses {8}. This finding indicates that conversation tends to prefer active voice and places a greater emphasis on actors – generally humans and animate referents – rather than happenings. In this, it is clearly distinct from more detached registers such as news scripts and formal prose.

Nominal morphology. Conversation had few occurrences of attributive adjectives {7} compared to more literary registers, but one of the highest levels of demonstrative adjectives {2}. Attributive adjectives are a means of integrating more information into the clause. Their lower occurrence level in conversation can be explained, as mentioned before, by the tendency of spontaneous spoken language to spread information out over multiple clauses and phrases. Demonstrative adjectives serve to refer back to a previous referent and bring it up within the context of a current utterance. They can also serve to distinguish between two or more similar referents in a discourse or temporal-spatial field. In a register with a greater scope, demonstratives are an important device.

In terms of its noun phrase complexity, conversation has a very small amount of NPs with more than 2 modifiers {8}. Once again, this finding is explained by the propensity of spontaneous spoken language to stage information in small chunks over a number of phrases or clauses. Noun phrase complexity was explored in depth in section 7.2. Here, it was seen that the conversational registers often employ extraclausal positions for complex noun phrases and has less of them in the subject position of transitive clauses. Conversation also has a small number of genitives relative {8} to other registers. This seems to be indicative of its less informational orientation. This interpretation is bolstered by the small number of total nouns in conversation {8}; the lowest of all of the registers. Finally, conversation had the highest grammatical concordance level of the spoken registers {4}. This is somewhat surprising, but might be seen to decrease with additional data. With a more varied sub-corpus, it would be possible to establish whether the particular sample of speakers used in this study were particularly conservative.

Prepositions, pronouns, and deixis. Related to the findings of the previous section is the small number of prepositions in conversation {8}, showing, once again, a tendency towards fragmentation and a less informational emphasis. Conversation has one of the highest levels of pronouns {2}, an indication of its need for referent tracking, yet it is perplexing that there is the large difference in the register between periphrastic possession with *aig* {1} and simple, pronominal possession {8}. At present, the way in which the discourse functions of the two constructions diverge is unknown and this is an area that

would clearly profit from additional research. Finally, conversation had the highest level of indefinite marking {1}, signifying a greater amount of vagueness and a reduced informational orientation.

Adverbs. Temporal adverbs are very common in conversation {1}. Amplifiers {6} and hedges {6} were not prevalent, which is out of line with previous research (Biber 1988) which found them to be characteristic of involved, oral discourse.

Lexical classes. Numerals were rare in the register {7}, indicating a low emphasis on information. Private verbs {3} were fairly common, a sign of involvement, while public verbs {6} and suasive verbs {8} were not. This indicates that the conversational texts in the corpus tended not to report the verbal actions of others nor indicate evidence of persuasive language.

English borrowing. There were a large number of unassimilated nouns {1} and non-nouns {1} indicating low formality. While there were a moderately high number of assimilated nouns {3}, there were less than registers that borrowed more self-consciously, adopting English words into Gaelic through phonological/orthographical change (i.e. formal prose, news scripts).

Lexical specificity. While the average letters per word in the texts were low {8}, conversation had the highest type token ratio of the spoken registers {5}.

8.2 Radio Interview

Information structure and discourse-level features. Radio interview was similar to conversation in its level of demonstrative clefts {2} and left-detachment {2}. There were no instances of right-detachment, but this may simply reflect the rarity of the feature rather than any register difference. Radio interview had a moderate level of fronted adverbials and peripherals {4} but there was not a statistically significant difference between it and conversation. However, if we consider the fact that it had a higher proportion of preposed subordination (.42) than conversation did (.33), the latter result becomes more salient.

Clause types. There were a relatively higher number of coordinate clauses {3} but a statistically significant smaller number of matrix/main clauses {5} than conversation. This difference may be due to the presence of an interviewer as well as the fact that turns are longer in radio interview. With more clauses, there are more opportunities for coordination. Radio interview evinced a moderately high level of relative clauses {3}; the highest for the spoken registers. It also had the highest incidence of cause/reason subordination {1}. This could be a reflection of the interview process, with questions begetting explanations. Conditional subordinate clauses were also frequent in the register {2}, perhaps an indication of greater hypothetical discourse.

Verbal morphosyntax. Progressive aspect showed up frequently in the register {2}, indicating a concentration on events that are on going within a particular tense orientation. Clausal negation was low {6} and significantly less than conversation {1}. Possibly, there is less overt disagreement in radio interview due to the workings of social mores, or perhaps it is reflective of a more informational orientation. There were high levels of copular verbs {2} and identificational clefts {2} but these results are difficult to interpret without additional research on their discourse functions. Direct interrogative clauses were frequently found {2} and imperative verbs were common as well {4} indicating a high level of interactiveness. Radio interview had the lowest return for past tense {8} pointing to very little evidence of narrative. On the other hand, it has the highest level of present tense {1}, demonstrating that it is concerned mainly with current

reference. As in most of the other spoken registers, there were no passive clauses with agents (0 occurrences) and few passive clauses overall {4}.

Nominal morphology. There were few cases of complex noun phrases {6} and a high amount of grammatical nonconcordance {2}, characteristics of spontaneous spoken language in general. The total number of nouns was low {7; cf. conversation at 8} indicating a low emphasis on information.

Possession, prepositions, and pronouns. Like conversation, radio interview had a high level of indefinite marking {2}, showing a tendency for vagueness. In addition, like conversation, it had a high level of periphrastic *aig* possession {2} whilst having a low level of pronominal possession {6}. As mentioned previously, it is difficult to interpret these results without further investigation. A relatively low level of prepositions obtained {6}, again a characteristic of less informative registers. However, there was a very high level of demonstrative pronouns {1}, suggestive of increased scope – reference to previous discourse and the pointing out of referents.

Adverbs. There was a high level of amplifiers {2}, emphatics {1}, and hedges {1} found, which are indicative of greater involvement and typically oral styles.

Lexical classes. Radio interview showed the highest amount of private verbs {1} in the corpus, denotative of greater involvement.

English borrowing. It had a high level of both unassimilated nouns {2} and unassimilated non-nouns {2}, both characteristic of informal, oral registers.

Lexical specificity. Overall, its average letters per word was low {6}, indicating fewer long words. Its type-token ratio was similarly low {6}, indicating a more narrow lexical range. These are both characteristic of spontaneous spoken language.

8.3 Sports Reportage

Information structure and discourse-level features. As a robust indication of the difference between the various clefts, sports reportage had the highest level of *'s e* clefts {1} whilst returning the lowest for *'s ann* clefts {8}. The *'s e* clefts in this register often have contrastive effect, such as in example (107):

- (107) *Nichol a' ruith ach 's e McStay a tha 'dol dha gabhail* [sic]
'Nichol running but it is McStay who is going to take it [the ball]'

Here, we are told that "Nichol [is] running" and we expect that he may get the ball. The pragmatic presupposition of the cleft is that someone will get the ball (probably Nichol), but we are told, in fact, that it is *McStay* who is going to get it.

Other clefts simply provide new information:

- (108) *'s tha e a-staigh tadhal ann an siud do Yugoslavia agus 's e Katernetch a fhuair an tadhal*
'and it is in a goal there for Yugoslavia and it is Katernetch who got the goal'

We are told that there has been a goal for Yugoslavia, which becomes the pragmatic presupposition of the following *'s e* cleft. The cleft constituent of that construction provides the new information, that it is *Katernetch* who got the goal. Quite often, clefts are employed in this way in the register: a particular thing is happening but we do not know who is associated with it, and then the cleft answers this question.

Sports reportage did not have a high level of demonstrative clefts {4} but they are still integral to the register, both in referring to previous discourse, as in

- (109) *agus sin leasan a dh'ionnsaich iad bhon a' ghèam an aghaidh na Cuimrich [sic] an-seo*
'and that is a lesson that they learnt from the game against the Welsh here'

or for pointing out referents in the temporal-spatial context:

- (110) *siud i tha i a-staigh tadhal do dh'Alba*
'that's it it's in a goal for Scotland'

Unlike other spoken registers, sports reporting did not have a high level of left-detachment {4} and showed no evidence of right-detachment. However, in this register, there is a dearth of finite verbs. The kind of role and reference strategies are much simpler than detachment; the most regular pattern is for a referent to be named as the RECEIVER or POSSESSOR of the ball in a prepositional phrase and subsequently, the subject of what looks to be a verbless clause in which we hear what action he takes. Ostensibly, the action is too rapid to use verbs in every clause and/or the ellipsis serves to make the message more urgent and exciting⁴², as illustrated by the following excerpt:

Sports reporting excerpt illustrating ellipsis:

am ball ma-thà aig Bett Bett le Johnson air an loidhne air an taobh a-muigh Bett fhèin a' ruith suas gu an deach Johnson a-steach dhan a' mheadhan gu McClair chaill McClair i tha i aig Hatsibegitch Hatsibegitch 'ga cur gu Bostardivitch Bostardivitch leatha a-mach gu Stanoykavitch Stanoykavitch air an loidhne air taobh thall na pàirc a-null dhan an taobh-sa ge-tà gu Spasitch Spasitch agus Nichol a' tighinn a-mach ma choinneamh

Translation:

the ball then at Bett Bett with Johnson on the line on the outside Bett himself running up until Johnson went in to the middle to McClair McClair lost it it's at Hatsibegitch Hatsibegitch putting it to Bostardivitch Bostardivitch with it out to Stanoykavitch Stanoykavitch on the line on the other side of the park over away to this side however to Spasitch Spasitch and Nichol coming out to face him

There were few fronted adverbials or peripheral elements {6} and a general preference of postposing subordinate clauses (.72 *vs.* .28 preposing) like most of the other registers.

Clause types. There was a high incidence of coordinate clauses {2}, which seems to be characteristic of monologic, spoken registers. As shown in §5.4.1, there was a tendency for coordination to be reserved for pivotal action sequences, such as ball stealing or advancing on a goal. There were a moderate number of main/matrix clauses {4}, but in interpreting this result, one must be aware that there was a high level of main verb ellipsis, as mentioned above; clauses were defined in the current research by the presence of a finite verb. There were very few instances of relative clauses {8}, cause or reason subordination {8}, concessive subordination {8}, conditional subordination {8}, and temporal subordination {8}. The overall impression is of a register that works in small chunks of information, with very little use of any kind of clause based modification.

Verbal morphosyntax. In terms of aspect, the one distinguishing feature in sports reportage was its high use of the progressive {1}, indicating its strong tie to on-going events. As seen above, many utterances in this register simply consist of a subject, a preposition or verbal noun – usually associated with progressive aspect – and an object.

⁴² Greenbaum and Nelson (1999) investigate various types of ellipsis in a number of different registers and find them replete in sports commentaries.

There was little clausal negation in the register {7}, characteristic of a more reportage based orientation. Sports reportage had the highest count of identificational clefts {1}, due to its need to specify referents but also due to the high number of clauses stating that 'it is' a throw-in, a bad play, an offside, a corner kick, or goal, etc. that is 'in it': e.g. *'s e sàdadh a-steach a bhios ann do dh'Alba* 'it is a throw-in that will be in it for Scotland'. There were very few interrogative clauses {7}, used exclusively between the commentators themselves. Similarly, there were no imperative verbs {8}. Both of these findings indicate very little interaction in this register. In terms of tense marking, there was a low return for the indefinite 2 {8} and a moderately low return for the past definite {6}. Both of these are associated with narrative based registers. On the other hand, there was quite a high occurrence of present tense {3}, which would have been higher if there had been more finite verbs used. It is possible in sports reportage to omit most finite auxiliaries (generally *tha*) because tense expression is superfluous – it is understood that all predicates, unless expressly marked as otherwise, refer to current time. Very few synthetic impersonals/passives occurred {8} and, indeed, very few passives overall {7}.

Nominal morphology. There were few demonstrative adjectives {7}, indicating a lowered demand on textual cohesion for distinguishing similar referents from one another. This may be an manifestation of the low scope in the register: the field of available referents is limited and well defined compared to a register like conversation. Sports reportage also had a low level of predicative adjectives {7}, which tends to be higher in more abstract or discursive registers. However, it had the highest level of appositives {4} and complex noun phrases {4} for the spoken registers. The genitive occurred with a low frequency {7} and the proportion of nonconcordant nouns to concordant ones was the highest of the register {1}. These are perhaps signs of the increased time pressure of its production circumstances.

Possession, prepositions and pronouns. Interestingly, this register had high levels of both periphrastic {2} and simple possession {2}. Many of the simple possession constructions were due to object incorporation in the aspectual particle *ag* (e.g. *McStay ga cur gu McClair* literally 'McStay at its putting to McClair') and the common reference to body parts, which are inalienable and usually take simple rather than periphrastic possession constructions. There were a large number of prepositions {3} and these were often associated with verbal nouns (see §6.3.2).

Adverbs. Sports reportage had the highest occurrence of place adverbs {1}, a sign of its high use of situation-dependent, exophoric reference. Time adverbs were also frequent {3}, denotative of a need for the placing of events in a temporal context. Amplifiers {1} and hedges {2} were both very common. This result shows the distinction in the corpus between reportage that is detached, i.e. news scripts, and that which is involved: spontaneous sports commentary.

Lexical classes. The only noteworthy finding from these features was that sports reportage had a very low number of public verbs {1} compared to the other registers. This clearly has to do with its emphasis on reporting action, not language.

English borrowing. Like many other kinds of broadcast registers, sports reportage had few unassimilated nouns {7} and non-nouns {7}. It also had few assimilated nouns {8}, but perhaps this had to do with the conscious attempt to use constructed phrases and neologism more than borrowing strategies (see §4.1.3.3).

Lexical specificity. It had the highest mean letters per word of the spoken registers {4}, due in part to the long surnames encountered in the texts. However, it had the lowest type-token ratio of the sub-corpora {8}, which may be related to its narrower scope.

8.4 Traditional Narrative

Information structure and discourse-level features. Traditional narrative had a moderately high occurrence of 's *ann* clefts {3} and demonstrative clefts {3} but few 's *e* clefts {7}. Demonstrative clefts have been characterised as helping to collect previous discourse, elucidate present concerns, and offer an entry into following discourse (Miller and Weinert 1998). It is not clear at this time what discourse function is conveyed by 's *ann* clefts. Traditional narrative also had a moderately high level of left-detachment {3}, a characteristic of spontaneous speech. Of all the registers, it had the greatest preference for preposed subordinate clauses (.48 *vs.* .51 postposed). This is partly attributable to the frequent appearance of temporal transitions between clauses.

Clause types. The level of coordination was very high {1} and it was suggested that this was to maintain the pace of the narrative and contain utterances into small, easily comprehensible chunks, providing an explicit demarcation of language flow. The level of cosubordinate clauses was also high [rank of 3 on mean; 2 on median], which seems to be a trait of narrative registers, but also of news scripts. A very high number of main/matrix clauses obtained {1}, indicating a preference for clauses with a small number of constituents. In general, there seemed to be a low level of subordination, but traditional narrative had the highest number of temporal subordinate clauses {1}, which seem to be used to place different episodes in relation with one another.

Verbal morphosyntax. Both incidence of perfect {7} and progressive aspect {7} were low. On the first count, it could be that there was a preference for perfective aspect, which is coded when the semantically main verb is in the first position (as opposed to being the verbal noun) and in the definite past (see 'Aspect' in Appendix 1). On the second, it may be that traditional narrative has a lower concern with on-going activities than completed ones. There were a large number of imperative verbs {1}, indicative of the amount of dialogue nested in the register. Past definite tense {1} occurred more frequently than in any other register, followed closely by the indefinite 2 or 'habitual-past' {2}. In contrast to this, there was a low return for present tense {7}. In terms of passives and impersonals, traditional narrative had one of the highest counts of synthetic constructions {2} but was not particularly marked as having a proclivity towards passive voice.

Nominal morphology. The low incidence of attributive adjectives {8} coupled with the few complex noun phrases {7} is characteristic of spontaneous spoken language in general, which paces descriptions out over multiple phrases and clauses. Traditional narrative had a very low instance of nonconcordant nouns {6}, and seems to be the most conservative, in this regard, of the spoken registers. The low number of nouns overall {6} is indicative of a less informational orientation.

Possession, prepositions and pronouns. There was a high return for pronominal possession {3}, but this result is difficult to interpret. The low number of total prepositions {7} is characteristic of a less informational orientation and a more fragmented structure, as prepositions are one of the ways of incorporating more information into the clause. There was a high count of demonstrative pronouns {2} which has correlated in previous research (Biber 1988) with greater involvement and orality. In addition, there was a very high return for pronouns in general {1}, showing a possibly greater demand on referent tracking over longer stretches of text.

Adverbs. A high count for place adverbs {2} obtained, a characteristic of spontaneous sports commentary and narrative, both of which place a high demand on them for situating purposes. There were low levels of amplifiers {7} and hedges {7}, showing a more detached orientation to the text.

Lexical classes. Traditional narrative had the lowest number of numerals {8} in the corpus, a sign of low informational orientation. However, it had the highest count of public verbs for the spoken registers {4}, a sign of greater speech reporting.

English borrowing. There were few instances of any lexical borrowing from English: assimilated nouns {7}; unassimilated nouns {8}; and unassimilated non-nouns {8}. §In view of this, the high rate of noun concordance, and the relatively frequent appearance of synthetic impersonals/passive, traditional narrative is one of the most linguistically conservative of the Gaelic registers.

Lexical specificity. The average letters per word {7} was low and the type token ratio {7} indicated a narrow range of vocabulary compared to most of the other registers.

8.5 Fiction

Information structure and discourse-level features. Fiction had the highest number of 's *ann* clefts {1} in the corpus. These are used to give emphasis to mainly prepositional phrases, but also relative clauses (marked sentence focus; see 1: §3.4.2.2), adjectives, and other non-nominal elements. However, as mentioned before, the ways in which its discourse functions contrast with the other kinds of clefts is not clear. There were a moderately high number of fronted adverbials and peripheral elements {3}, indicating that this kind of guideposting is important for the register, especially for establishing time, place, and manner (see examples in §5.3).

Clause types. The highest level of cosubordination was encountered in fiction {1}. Main/matrix clauses occurred with high frequency {3}, indicating that clauses tended to be small. There were few complement clauses {8}, indicating greater editing and a less informational orientation. Both concessive {1} and conditional clauses {1} occurred with high frequency, being more characteristic of imaginative or less reportage-based registers. Temporal clauses {2} also returned with a high frequency; there was need to place predicates in temporal relation with one another.

Verbal morphosyntax. Clausal negation was high in this register {2}, a feature which seems to correlate with less fact-based registers. The high return for imperative verbs {2} and direct interrogative clauses {3} is denotative of the amount of dialogue represented in the register. There was a high incidence of both the indefinite 2 ('habitual-past') {1} and past tense {2}, both symptomatic of narrative. In contrast, the level of present tense {8} was the lowest in the corpus. Passive clauses obtained with the lowest frequency for the written texts {6}, suggestive of a greater emphasis on actors rather than events, recipients, undergoers, etc.

Nominal morphology. Fiction had the highest level of attributive adjectives {1}, which occur with greater frequency in more edited texts. The level of complex noun phrases however {4} was not overly marked. Although a small number of genitives obtained relative to other written registers {4}, fiction had one of the highest indices of noun concordance {7}.

Possession, prepositions and pronouns. Pronominal possession occurred with the greatest frequency in this register {1} but this is difficult to interpret without further research. Pronoun counts were moderately high {3}, indicative of the need to track multiple referents over extended discourse.

Adverbs. No marked features.

Lexical classes. Numerals {6} were not frequent in fiction, evidence of its low emphasis on information. Private verbs counts were high {2}, but also public verbs {2}, reflecting the presence of speech reportage wherein speakers talked about themselves and opinionated.

English borrowing. Fiction tended to have a moderate amount of borrowing: assimilated nouns {5}; unassimilated nouns {4}; unassimilated non-nouns {5}.

Lexical specificity. Although the words in fiction were of middle length {5}, relative to the other registers in the corpus, there was a high type-token ratio {2}, indicating a varied lexicon.

8.6 Formal Prose

Information structure and discourse-level features. Like fiction, formal prose had a high number of 's *ann* clefts {2}, but the discourse function of this feature remains unclear. There were very few occurrences of left-detachment {6} and none of right-detachment, denotative of formal, edited prose. Guideposting in the form of fronted adverbials and peripheral elements was more common than in any other register {1}.

Clause types. There were few coordinate clauses {7}, showing a higher degree of integration. Also, there were few main/matrix clauses {7}, suggestive of longer and fewer clauses overall. Relative clauses occurred with the greatest frequency in formal prose {1}, again suggesting a greater amount of editing and integration. Concessive clauses also occurred at a relatively high level {2}, seemingly a sign of more abstract or non-reportage based registers.

Verbal morphosyntax. Progressive aspect obtained the most infrequently in formal prose {8}. This may be related to a greater emphasis on abstractions rather than current matters. Both interrogative clauses {6} and imperative verbs {6}, were rare, an indication of a low interaction level. Present tense obtained less than most other registers {6}, again symptomatic of a focus on non-immediate context. In terms of valence decreasing expressions, formal prose had the highest level of synthetic impersonals/passives {1} and a high level overall of total passive clauses {2}. These are signs of a more detached, more formal orientation.

Nominal morphology. Formal prose had the second highest level of attributive adjectives {2}, a clear sign of integration in the register. Both demonstrative adjectives {1} and predicative adjectives {2} also obtained frequently. Demonstrative adjectives are involved in textual cohesion and in distinguishing referents from one another. Predicative adjectives have been weakly correlated with more abstract registers (Biber 1988).

The occurrence of complex noun phrases {1} is higher in formal prose than the other registers, again indicating greater integration and editing. The examples in §7.2 showed the variety and complexity of NP modification in the register. A major difference between formal prose and the conversational registers was its less frequent use of extracausal positions with complex NPs. Genitive nouns {2} occur with a high frequency along with nouns in general {3}, both signs of a higher information emphasis. Bolstering this impression, there are very few instances of indefinite marking {8}. Finally, formal prose exhibits the most conservative nominal morphology, as measured by its low incidence of nonconcordant nouns {8}.

Possession, prepositions and pronouns. There was a low level of periphrastic possession {7} in the register but the discourse motivations of the construction need further clarification before it can be interpreted. The high frequency of prepositions {1} supports the view of

formal prose as having a high information emphasis. The low level of pronouns {7}, on the other hand, could be reflective of a more detached, less narrational perspective.

Adverbs. Both instances of place adverbs {7} and time adverbs {8} are very low in this register. This suggests that it is more abstract and less situation-dependent.

Lexical classes. Formal prose had one of the highest occurrences of numerals {2}, again denoting a more information-based register. The incidence of both private {8} and public {7} verbs was low, indicative of greater abstractness.

English borrowing. Formal prose had the highest level of assimilated nouns {1}, which shows a conscious attempt to dress most English words in Gaelic orthography. It had a moderately low level of unassimilated nouns {6} and unassimilated non-nouns, indicative of greater formality and editing.

Lexical specificity. The high average letters per word {2} indicates a greater number of long words. However, formal prose had the lowest type-token ratio of the written registers {4}, which is somewhat surprising, but it could be due to the greater repetition of terms in the register, as found by Biber *et al.* (1999) in academic prose.

8.7 News Scripts

Information structure and discourse-level features. News scripts had the lowest returns for two kinds of clefts – 's e {8} and demonstrative clefts {8} – as well as having a very low frequency of 's ann clefts {8}. The lack of these features seems to indicate low scope in the register, for instance, in returning to previous discourse, summing up, or contrasting points of reference. The absence of left- {8} or right-detachment {8} is denotative of a formal, written register and a high level of integration. Perhaps another indication of the straight discourse movement of news scripts is the low occurrence of fronted adverbials or peripherals {8} and its overriding preference for postposed (.93) rather than preposed (.07) subordinate clauses. Overall, there is little adverbial guideposting of the kind to which Chafe (1982) referred.

Clause types. The low amount of coordinate clauses {8} obtained in this register, again indicates little continuation of topic as well as higher integration. However, there was a high level of cosubordinate clauses {2}, shared with the narrative-based registers. More research needs to be on the discourse properties of this construction. News scripts had the lowest number of main/matrix clauses {8}, which indicates longer clauses and perhaps more clause combinations. Relative clauses {6} were less frequent than other kinds of written registers, in addition to subordinate clauses of cause and reason. Complement clauses {1} returned with a high frequency indicating a high informational orientation. Temporal subordination {7} was infrequently found relative to the other registers, indicating that news scripts did not have a need for placing different propositions in temporal relation with each other.

Verbal morphosyntax. Perfect aspect {1} was most common in news scripts. This could be a product, as previously mentioned, of translation from English, or be indicative of reportage styles which deal with events that occurred in the past but are still relevant in the current time frame. Progressive aspect {3} also occurred with a moderately high frequency, indicating a concern with on-going events. News scripts had the lowest count of clausal negation {8}, which denotes a more factual, report based orientation. The number of copular verbs {8} and identificational clefts {8} was similarly low, but more research needs to be done on their discourse properties before these results can be interpreted. Consistent with its more detached orientation, low levels of interrogative

clauses {8} and imperative verbs {7} were found. There was a low occurrence level of both past tense {7} and the indefinite 2, indicating a style that places emphasis on current, on-going events and that is less concerned with hypotheses or habitual occurrences. Both agent-specified passive clauses {1} and passive clauses overall {1} occurred with the highest level in news scripts, again characteristic of a more detached orientation. However, there was a lower level of synthetic impersonals and passives {5}, suggestive of a less conservative morphology.

Nominal morphology. Attributive adjectives {3} occurred moderately frequently in this register, a characteristic of more informational and integrated text. In contrast, there were few demonstrative {8} or predicative adjectives {8}. Demonstrative adjectives act as textual cohesion devices and appear to occur with greater frequency in more discursive registers such as formal prose and conversation. They would not be as heavily indicated in texts with quick topic changes and little back discourse movement. The low incidence of predicative adjectives seems to indicate a style that is more factual and less abstract. Appositives {1} occurred with a much greater frequency than in the other registers and seems to be the tool of preference in news scripts for specifying referents. Complex NPs {2} and genitive nouns {1} were frequently encountered, correlates of the register's more informational and integrated style. However, news scripts had one of the highest rates of nonconcordant nouns {3} and nonconcordant adjectives {2}, perhaps a sign of its stringent production circumstances. The high overall noun count {1} is a token of their high informational orientation.

Possession, prepositions and pronouns. The few cases of indefinite marking {7} specify a less vague, more informational style. Perhaps denoting greater detachment overall, there were low returns for both periphrastic {8} and pronominal possession {7}, but more research needs to be done on the discourse functions of these features. The high total number of prepositions {2} indicates an informational style, like several of the other features mentioned above. There were few demonstrative pronouns {8}, indicating less back discourse movement and symptomatic of the frequent changes of subject in the scripts. Finally, there were few pronouns in the scripts {8}, showing the scant need in this register for managing multiple referents.

Adverbs. The low count for place adverbs {8} is indicative of a low concern with the explicit establishment of location in news scripts. On the other hand, it had one of the highest occurrences of time adverbs {2}. Both amplifiers {8} and hedges {8} were infrequent, suggesting a more detached orientation.

Lexical classes. Numerals {1} were common, again a sign of the informational style of news scripts. While private verbs {7} were rare, public verbs {1} were common and indicative of the amount of reported speech found within the texts. Suasive verbs {1} were also common and it was suggested that there is a by-product of the large amount of persuasive speech, from individuals such as politicians, that is reported in the news scripts.

English borrowing. Assimilated English nouns {2} were frequent. Previous research (Lamb 1999) has found that news script translation has become more lexically oriented than it was in the past and this feature seems to bolster this interpretation. Unassimilated nouns {3} and unassimilated non-nouns {4} were both moderately frequent in the register as well.

Lexical specificity. News scripts were observed to have the highest average letters per word {1} and highest type token ratio {1}, thus it uses the longest words and has the greatest vocabulary range of the corpus.

8.8 Popular Writing

There were very few features on their own that distinguished popular writing from the other registers. The results indicate that it occupies a middle position between the kinds of axes or dimensions used to characterise the differences between speech and writing, formality and casualness, involvement and detachment, and so.

Information structure and discourse-level features. Demonstrative clefts {7} were infrequently found in popular writing, perhaps a sign of little returning to previous discourse. Both left- {7} and right-detachment (1 occurrence in 10,000 words) were rare, indicating a register with less oral characteristics and more editing. Signifying greater guideposting and transitions with previous discourse, there was a high return for fronted adverbials and peripheral elements {2}.

Clause types. Indicated by the lower number of main/matrix clauses {6}, clauses in popular writing tend to be longer than in spoken registers, but shorter than in other kinds of writing besides fiction. Relative clauses {2} were common, and indicate a greater level of editing and integration. The level of complement clauses {2} seems to reflect a tendency towards a more informational orientation in this register.

Verbal morphosyntax. Perfect aspect {2} occurred with a high frequency, a tendency of more literary registers. The results indicated that popular writing differed significantly from more reportage-based registers in terms of clausal negation {3}. A moderately high level of total passive clauses {3} obtained, with statistical tests indicating that popular writing was more akin to the detached registers such as news scripts and formal prose in this regard.

Nominal morphology. The high level of predicative adjectives {1} may indicate a more abstract than factually orientated register. Appositives {2} occur with a high frequency in this register and tend to be used to supply additional information about, or name referents as in example (111):

- (111) *fhuaire na beachdan seo taic bho fhear labhairt Gàidhlig nan Nàiseantach, Micheal Ruiseal, ...*
 'these opinions received support from the Nationalist Gaelic spokesman, Michael Russell, ...'

Complex NPs {3}, genitive nouns {3}, and nouns overall {2} are frequent in popular writing, and indicate a more informational orientation and greater level of integration. There was a fairly low proportion of nonconcordant to concordant nouns {5}, indicating greater editing and morphological conservativeness.

Possession, prepositions and pronouns. The high level of prepositions {3} indicated more clausal integration and a more informational orientation than the spoken registers. There was a low occurrence level for demonstrative pronouns {7}, suggestive of less discourse motion than in most other registers.

Adverbs. The low incidence of place adverbs {6} and time adverbs {7} is a sign of the register's lower reliance on time and physical space for the grounding of discourse. That is, in this regard it seems to have a more abstract orientation, similar to formal prose.

Lexical classes. No marked features.

English borrowing. Popular writing was not particularly marked in its use of English-origin lexemes. In general, it used fewer unassimilated items than spoken registers, but more than the more lexically conservative written registers such as news scripts and formal prose.

Lexical specificity. The average letters per word {3} in popular writing tended to be longer than most other registers, as did its breadth of vocabulary, measured by the type-token ratio {3}.

9 Conclusions

9.1 Summary

The purpose of this study is to examine the way in which grammar and context interact in a sample of eight different textual varieties of Scottish Gaelic. These varieties of text were selected to be representative of a wide range of both linguistic and contextual variation. The primary theoretical assumption underlying this work is that linguistic behaviour in repetitive contexts tends to become crystallised to a detectable extent. In other words, the linguistic forms produced by humans reflect, in part, the situational parameters and message types inherent to the context in which they are produced. The pairing of language and context is formalised in the term 'register'; this is stated thus, because it is also undeniable that language itself contributes to context. Another basic assumption of the research is that co-occurring linguistic features in a given context can be interpreted as sharing one or more communicative functions. In order to determine these functions, it is necessary to examine the different contextual parameters present in the corpus, and consider the ways in which associations between these and linguistic form are conventionalised. Also, as has been clear from past research, it is necessary to consider how planning time and the limits of human cognition – especially short-term memory – interact to produce certain largely non-volitional aspects of spontaneous spoken language.

From an original tag set of nearly one hundred discrete linguistic features, 63 were selected for inclusion in the study based upon frequency and salience. The distribution of these features, which were compared statistically across the 8 registers, uncovered a wide range of linguistic variation in the corpus. A profile of each register was compiled based upon its particular feature distribution (see Chapter 8). These profiles were consulted for evidence of feature groupings that appeared to have similar functional attributes. In addition, the feature distributions themselves, across the 8 registers, were examined in the likelihood that adjacent, outlying registers would use a feature for similar purposes. Finally, a thorough quantitative and qualitative study was made of nominal concordance and complex nouns, two features that evinced robust differences between the registers in the statistical analysis. As a result of these investigations, a number of observations were made, as follow below. For the moment, these will be framed mainly in terms of register contrasts rather than in terms of underlying contextual influences, which will be addressed in the following section.

1. Scottish Gaelic conversation (represented in this corpus by C and RI) is significantly different from other types of registers in its lexis, morphology, and syntax. Lexically, there tends to be more English borrowing, more indefinite marking, fewer complex noun phrases, fewer nouns in general, fewer genitives, fewer attributive adjectives, and smaller words. Morphologically, one finds generally more nonconcordant nominal forms. Syntactically, structures are simpler and evince the gradual build-up of information across many clauses discussed by Miller and Weinert (1998) and Chafe (1982; and Danielewicz 1987). This characteristic is manifested in more (hence smaller) main clauses, fewer cases of certain subordinate clauses, and more fragmented structures, such as detachment and the asyndetic juxtaposition of clauses and phrases. Also in evidence are features linked to interaction: conversation has the most interrogative clauses and the third most imperative verbs in the corpus (behind traditional narrative and fiction).
2. Almost diametrically opposed to conversation is formal prose. In addition to the opposite degrees of the above characteristics, one also finds a greater proportion of passives and impersonals – especially inflected ones – which is characteristic of a more 'detached' style. In general, syntactic relations in formal prose are more explicit and meaning is built up through a greater specificity of nominal forms than in

conversation. This is reflected in the high noun counts and higher average word length in formal prose.

3. There were robust differences between narrative and other register types in the corpus. Narrative showed more past tense, more of the indefinite 2 ('past-habitual'), more pronouns and less present tense. There were also more coordinate and cosubordinate clauses. Finally, time reference in these registers was more endophoric, indicated by the high concentration of temporal clauses. As opposed to news scripts, in which time reference is presented as static, time in narrative unfolds, built up and stabilised through chains of propositions and the explicit relations which are coded in temporal subordinators.
4. Interactive registers, such as conversation and radio interview, tended to differ from those with a more informational orientation, such as formal prose and news scripts. In general, interactive registers show more private verbs, more amplifiers and hedges, more clausal negation and generally more present tense, indicating a greater salience of current time frames. Informational, detached texts were characterised by an absence of these features as well as the presence of others such as passives, numerals, higher type-token ratios, longer words, more attributive adjectives, and prepositions. These indicate that these registers place a greater emphasis on the conveying of information than the actions of human subjects and their beliefs, opinions, perceptions, etc.
5. Fact-based, reportage registers (news scripts and sports reportage) differed in specific ways from most of the other registers in the study. The main distinction was a lowered proportion of features serving to relate back to, or make reference to, aforementioned clauses and constituents: demonstratives; temporal subordination; preposed subordinate clauses; and fronted adverbials or peripheral elements. It was suggested that this was due to the strong forward inertia present in these registers, with frequent topic changes and little discursiveness in evidence.
6. There were major differences in grammatical conservativeness evinced between the registers. In general, the more typically oral, more spontaneous registers had more grammatical nonconcordance while more literal, more edited registers had less. However, news scripts, which are written, had one of the highest rates of nonconcordance while traditional narrative, which is spoken, had amongst the lowest. There was also a difference between the types of passive forms found in the more 'formal' registers: while traditional narrative and formal prose had a high proportion of inflected passives, news scripts tended to prefer periphrastically-formed ones. These findings are somewhat paradoxical and will resurface below where the contributing factors of context are explored further.
7. As mentioned above in terms of news scripts and traditional narrative, the registers differed in the ways that they made reference to time. There were two main contrasts: a) registers had a high proportion of markers of current reference (i.e. present tense and progressive aspect) or did not; and b) registers preferred either endophoric or exophoric temporal reference, with the former manifested in more temporal subordination and the latter in more temporal adverbs. The non-narrative spoken registers (i.e. conversation, radio interview, and sports reportage) had the highest occurrence of current reference while the narrative registers had the lowest. Endophoric temporal reference was most characteristic of the narrative registers while exophoric reference was more common in news scripts.

In the following section, conclusions will be made on the conventional associations between context and linguistic form in the corpus. After this point, it will be appropriate

to consider the findings of past register research in light of the current study. Its methodology differs from that devised by Biber, but it is worth examining whether the same kinds of contrasts that he detected through factor analysis are indicated here. Following this discussion, concluding remarks will be made about the grammar of spontaneous spoken Gaelic. In anticipation, it will be maintained that conversational Gaelic, in particular, differs linguistically in a number of basic ways from other registers. In addition, there are variations between the different forms of Gaelic conversation – e.g. public interview and private conversation – that were examined in this study. The findings of this research, especially the aforementioned, have implications for future studies on Scottish Gaelic grammar and register, as well as for pedagogy and resource development. In the final section, these will be highlighted and suggestions for further work will be made.

9.2 Associations Between Context and Linguistic Form

Context was defined broadly, encompassing the aspects of situation detailed in Chapter IV. Of these, the results indicate the involvement of five main contextual features⁴³ in the variation of the eight Scottish Gaelic registers studied here: production constraints; discourse freedom; information orientation; interaction; and producer characteristics. The following discussion describes the associations between these contextual features and the various linguistic forms surveyed in the corpus. Following this section, the results as a whole will be considered in light of past register studies to consider how well they reflect the conclusions made by authors working in languages other than Gaelic.

Factors related to production constraints exert the most powerful influences on the linguistic forms of the registers surveyed for the study. The most constrained of the registers is sports reportage, where the addresser attempts to convey every significant movement of the players and ball in real time. There is almost no preparation time possible as events must be conveyed in immediate succession to their occurrence. Consequently, utterances appear stripped of any features not crucial to the message at hand. The clearest manifestation of this is the high rate of main verb ellipsis in the register (see discussion of Sports Reporting features, beginning on page 176). Conversation and radio interview are less constrained than sports reportage, but more so than all of the written registers surveyed here. As a result, language in conversation is fragmented in form, with information distributed gradually across multiple clauses and phrases, as seen in the greater number of clauses overall and constructions such as detachment (§5.2) and clausal juxtaposition (§5.4.5). Also attributable to lower planning time, complex noun phrases are less frequent in conversation, as covered in sections 6.2.3 and 7.2, and there is less lexical variety (§6.6), generally fewer relative clauses (§5.4.4.2) and more grammatical nonconcordance present (§s 6.2.4.2 and 7.1) than in the written registers. At least partly due to production constraints is the greater proportion of English loans in Gaelic conversation. However, there was a difference between radio interview and private conversation in this regard (see §6.5.3), which was attributed to the emergence of self-consciousness in the face of greater accountability in the public register. Like the conversational registers, traditional narrative also shares certain morphosyntactic features related to plannedness, such as left-detachment, a high number of main clauses, and few instances of complex NPs, but it differs significantly from the other spoken registers in terms of its grammatical concordance and level of English borrowing. Contextually, traditional narrative is less typically oral than the other spoken registers in that it is

⁴³ To these, narrative could be added, but at the risk of tautology: narrative, in and of itself—as Biber (1994) points out—is not a situationally well-specified register. Although it is clear that narrative discourse differs linguistically from other kinds of discourse, as detailed in the previous section, to say that narrative diverges from other register types because it is narrative is to succumb to circular reasoning. Probably, the contextual features contributing to this difference are narrative's orientation to past events, its low informativity, its relaxed production constraints in the case of fiction—permitting copious editing and opportunities for syntactic integration, lexical diversity, etc.—its concern with primarily sequential, dynamic events rather than static ones, and the significance given to the behaviour, cognition, and linguistic exchanges of specific human (or anthropomorphic) participants.

rehearsed and the product of many speakers over time. These factors, and the influence of producer characteristics (see below), contribute to its lower ungrammaticality and English borrowing. Of the written registers included in the corpus, only news is affected by production constraints. Although in many ways, it is a typically literate register (e.g. more complex NPs, genitives, numerals, a high type token ratio, fewer 'fragmented' features, etc.), it evinced a high return for nonconcordant nouns, which is partly a sign of the lowered planning time available and the reduced amount of editing compared to the other written registers surveyed.

Discourse freedom concerns the degree to which the navigation of earlier points of reference is inhibited in a register. In the two reportage registers included – news scripts and sports reportage – addressers are bound to continually advance the discussion, with little returning to the content of previous utterances. Correlating with such an orientation, described as 'strong forward inertia', is the paucity of certain linguistic features functioning to associate blocks of language with one another and re-activate earlier points of reference. So-called guideposting constructions, such as preposed subordinate clauses and fronted adverbials/peripheral modifiers (see §5.3), are less frequent in these registers. These manage transitions and explicitly mark relations between different clauses. Demonstrative adjectives are also rare (§6.2.1), serving to refer anaphorically to previously mentioned referents. In sports reportage, over half of the demonstrative adjectives in evidence occurred at lapses in the game – during commentary – bolstering the impression that they are more often found in discursive language.

The degree of each register's **informationality** is reflected in particular linguistic characteristics. Formal prose and news scripts are the most overtly informational of the registers. These contrast to registers such as conversation, traditional narrative, and fiction, which are more oriented towards the interactions of human participants and describing their mental states and speech. In formal prose and news scripts, there were many features that have been linked in past research with the conveyance of information. Informational text tends to be nominally rich and specific, denoted in the registers by a high occurrence of nouns (§6.2.4.1), genitives (§6.2.4.3), complex NPs (§6.2.3) and long words (§6.6), and fewer cases of indefinite marking (§6.2.4.4). Prepositions (§6.3.2), which Chafe (1982; and Danielewicz 1987) describes as serving to integrate information within idea units, are encountered with the highest frequency in these registers, as are passive/impersonal constructions (§6.1.3). The latter give prominence to actions and objects while de-emphasising usually human agents. Informationally speaking, reference to agents in such texts is often superfluous and/or can be retrieved from context.

Interactional discourse, such as in conversation and radio interview (but also nested within fiction and traditional narrative), tends to differ in several ways from monologue and single-addresser writing. Superficially, there are more interrogative clauses and imperative verbs (see §6.1.5), as one would expect. There is also a greater number of private verbs (§6.5.2), which communicate the mental states and processes of participants. Finally, the instance of clausal coordination (§5.4.1) seems to decrease in spoken text as the number of participants increases. Spoken monologue (i.e. traditional narrative, followed by sports reportage) had the greatest amount of coordination, and the conversation text with the highest proportion of coordination also had the lowest number of participants.

Finally, the influence of **producer characteristics** has the most tangible effects on the differences between traditional narrative and news scripts. Although these registers are from different modes, their production characteristics gravitate them towards one other. Traditional narrative is a less typically oral register than the others in the study due to the fact that it is rehearsed and the product of many authors over time. Similarly, the news script register is less typically literate due to the time constraints built into its context. It was seen that traditional narrative significantly differs from news in the distribution of certain features which have been linked to grammatical conservativeness:

the proportion of nominal concordance; English loans; and synthetic passive/impersonal forms (see references in §2.4). One would expect, given its mode and arguably greater formality, that news language would be the more conservative, but this is not the case. In explaining this discrepancy, it is useful to examine the producers themselves. The clearest difference between the two sets of producers is age. The reciters of the narratives used in the corpus were at least a generation, if not two or three, older than those who were working in the BBC Gaelic department 11 years ago, when the texts were collected. As Gaelic is a dying language, it follows that those born earlier will have had access to a richer, less English-dominant linguistic experience. These earlier-born speakers, although usually bilingual, were often Gaelic dominant. Furthermore, as caretakers of what they viewed as important, cultural material, they were taking an interest in and preserving older forms of the language. In sum, these speakers had an exposure to, facility with, and interest in more 'conservative' Gaelic. In contrast, the writers and deliverers of Gaelic news today are younger, generally less Gaelic dominant⁴, in contact with more English speakers, and living at a time when Gaelic is not necessary for most the world outwith their jobs. Consequently, they have access to a more restricted reserve of lexicon and idiom than the producers of traditional narrative, as well as less facility with certain aspects of the language's morphology. Coupled with the influence of translation effects, this explains the fewer synthetic impersonals present in news scripts *versus* traditional narrative, and the greater number of nonconcordant forms and English loans.

9.3 A Review of the Findings in Terms of Past Register Research

As detailed in Chapter II, the earliest register studies as such expressly set out to explore the differences between the modes (e.g. Blankenship 1962; O'Donnell 1974; Poole and Field 1976). It was pointed out that these were problematic because of issues relating to the elevation of certain registers over others as representative of speech or writing. Modern studies have moved away from an emphasis on modal differences towards a more fine-grained approach based upon register, generally accepting Biber's (1988) conclusion that there are no definitive differences between spoken and written English. He found that there are no linguistic features whose distribution divides all spoken registers from all written registers. In the current study, which used far fewer registers than were included in Biber's study, the following features were found to have apparently modal distributions:

SPEECH

demonstrative clefts
left detachment

WRITING

perfect aspect
attributive adjectives
complex noun phrases
genitives
lexical variety

However, for all of these features apart from demonstrative clefts and perfect aspect, the difference can be ascribed to production constraints (i.e. planning time), as covered in the previous section. The higher occurrence of demonstrative clefts in speech is probably due to its generally greater discourse freedom, and the higher occurrence of perfect aspect in writing may be due, in part, to translation effects (see §6.1.1). Considering the fact that, of the 63 features that this study investigated, only 7 were distributed modally, Biber's statement is supported for this data overall.

Relating to plannedness, many of Ochs' (1979) main findings were replicated here. Her finding that semantic relations between referents and following predications were not explicitly marked (i.e. in detachment constructions) in unplanned texts, is also a characteristic of the Scottish Gaelic data. This feature was linked to the general tendency

⁴ Since the 1990s, there have been more cases of semi-fluent speakers delivering the Gaelic news, echoing the situation in Ireland.

of spontaneous speech to distribute information gradually across many clauses and phrases, a manifestation of low planning time, the limits of human short-term memory, and perhaps an awareness on the part of addressers of their listeners' comprehension limits. Ochs indicated that more planned, literate texts evinced more explicit marking of joined propositions than more characteristically 'oral' texts, which evinced asyndetic clausal juxtaposition. The current study similarly found that there were generally fewer subordinate clauses in the spoken registers – particularly relative clauses – and that juxtaposition was an important strategy for relating propositions in speech, providing greater syntactic flexibility than subordination. Ochs' finding that meaning in planned texts tended to be more lexicalised was also detected in the current study. In general, more planned and edited Gaelic registers tended to have a greater concentration of nominal features, greater word length, and greater lexical variety. However, one finding that was not replicated in the current study, detected by both Ochs and Tannen (1982b), was the tendency for spoken narrative to evince the 'historical present'. In fact, traditional narrative had virtually the same occurrence level of present tense as written fiction (25.24 *vs.* 24.88 occurrences/1000 words). This highlights a basic stylistic difference between Scottish Gaelic and English.

The two dimensions that Chafe (1982) posed as basic to the difference of speech and writing – integration *vs.* fragmentation and detachment *vs.* involvement – are borne out by the Scottish Gaelic data. Most of the properties mentioned in results 1 and 2 above, at the beginning of section 9.1, concern the former distinction. There was also a tendency for Gaelic speech to have more interactional/involved features than writing, such as private verbs, interrogatives, indefinite marking⁴⁵, as well as fewer passives. Overall, Chafe's dimensions fit the data, but they do not account for all of the variation found here. Some of the factors important to this study but not covered by these dimensions are discourse freedom, informationality, narrative discourse, and temporal orientation.

The results of Biber (1988) and others working within the multidimensional framework take account of a much greater range of variation than previous investigations. Although the current study is not multidimensional, it is possible to compare the underlying bases of variation detected here with the factors presented in these studies. Biber (1988) found six main dimensions of variation, the first three of which are salient to the present data: Informational *vs.* Involved Production; Narrative *vs.* Non-narrative concerns; and Explicit *vs.* Situation Dependent Reference. He avers that his first factor powerfully summarises a basic aspect of language variation, between an informational, nominal style and one that is involved and verbal. In effect, this dimension condenses both of Chafe's (1982) into one. The main linguistic features loading on the dimension that were also employed in the present research are as follows:

Table 119: Biber's (1988) Dimension One: Informational *vs.* Involved Production

<i>Positive: verbal style</i>	<i>Negative: nominal style</i>
private verbs	nouns
present tense verbs	word length
clausal (analytic) negation	prepositions
demonstrative pronouns	type-token ratio
causative subordination	attributive adjectives
indefinite marking	
hedges	
amplifiers	
(WH) questions	

⁴⁵ This is an interactional feature as it calls upon the addressee to fill in the gap posed by the indefinite marker.

Biber mentions that some of the most positively loading registers on Dimension 1 are conversation and interview while some of the most negatively loading ones are academic prose and press. For the Gaelic data, this dimension almost parallels the differences that were found between conversation and formal prose. The linguistic features in evidence and the way that the registers loaded upon it are also very similar to the first dimensions of both Kim (1990) and Jang (1998). Other linguistic features in Scottish Gaelic that correlate with more 'nominal' styles are relative clauses, assimilated borrowing, complex noun phrases, genitives, and passives. Features correlating with a 'verbal' style in Gaelic include detaching constructions, demonstrative clefts, nonconcordant nominal forms, a greater number of main clauses, and interrogative clauses. In sum, the basis of Biber's Dimension 1 seems well substantiated in the Gaelic data.

Biber's Dimension 2 relates specifically to narrative *vs.* non-narrative discourse. The linguistic features related to this contrast which were also employed in the present study are as follows:

Table 120: Biber's (1988) Dimension Two: Narrative *vs.* Non-narrative Concerns

<i>Positive</i>	<i>Negative (all weak correlations)</i>
past tense verbs	present tense
3 rd person pronouns	attributive adjectives
perfect aspect	word length
public verbs	

Again, this contrast is reflected in the Gaelic data as well, as pointed out in finding 3 of §9.1. Narrative in Gaelic also tends to evince a greater proportion of the indefinite 2 ('past-habitual'), more coordinate clauses, and more cosubordinate clause. Time reference seems to be marked in Gaelic narrative as well, as there is a higher level of temporal clauses than in other register types. Jang and Kim also found a contrast between narrative and non-narrative discourse, although the linguistic features loading on the factor in these cases were different. Interestingly, for both Taiwanese and Korean, time adverbs were a significant feature of narrative, which is not reflected in the Gaelic data. However, the fact that all four languages have found significant differences in narrative *versus* non-narrative discourse indicates that it may be a basic dimension of human language register variation, as claimed by Biber (1995).

Biber's Dimension 3 is only partly reflected in the Gaelic data. He interprets this dimension as signifying "highly explicit and elaborated, endophoric reference from situation-dependent, exophoric reference" (p. 142). Some of the linguistic features comprising the factor in Biber (1988) are below:

Table 121: Biber's (1988) Dimension Three: Explicit *vs.* Situation Dependent Reference

<i>Positive</i>	<i>Negative (all weak correlations)</i>
relative clauses	time adverbs
pied piping	place adverbs
phrasal coordination	adverbs

If this had been a multi-dimensional study, perhaps it would have located a similar dimension, but the only evidence that can be advanced for it given the present results, is the distinction between the primarily endophoric temporal reference of traditional narrative, and the primarily exophoric temporal reference of news scripts. Biber's results indicate that sports reportage loads lower on this dimension than any other register, being overwhelmingly situation dependent. Kim's (1990) dimension 5 ('On-line Reportage of Events') was also typified, at one pole, by sports broadcasts. In the Gaelic data, there were

a few robust markers of sports reportage: progressive aspect; place adverbs; few cases of temporal subordination; and few cases of the indefinite 2. However, it does not seem that the dimensions found by Biber and Kim agree in total with the tendencies in the Gaelic data.

The rest of Biber's (1988) dimensions, as follows, are not transparently applicable to the Gaelic data: Overt Expression of Persuasion; Abstract *vs.* Non-abstract Information; On-line Informational Elaboration. However, if we view the results in terms of Biber's (1995) 'universal' dimensions of register variation, summing up the findings of previous multi-dimensional studies on a small number of unrelated languages, some of them seem substantiated in the present study. The first 'universal' dimension – Type A – concerns a range between interactive and information-oriented discourse. This is similar to the Dimension 1s of Biber (1988), Kim (1990), and Jang (1998) and the 'involved *versus* detached' contrast mentioned by Chafe (1982). The second universal dimension – Type B – deals with production circumstances, typically characteristic of registers such as conversation and sports reportage, where production time is maximally constrained. This dimension type seems conflated with universal Type A in Biber's (1988) Dimension 1, but also shows up in his Dimension 3. Kim's (1990) Dimension 5 ('On-line Reportage of Events'), essentially marks this contrast and it recalls the 'fragmented *versus* integrated' distinction made by Chafe (1982). Production circumstances clearly have a pervasive impact on linguistic form in Scottish Gaelic, as was discussed in §9.2. Universal Type C – a dimension related to personal stance – is only somewhat substantiated in the Scottish Gaelic. The linguistic features that Biber (1995) mentions as being characteristic of this dimension are hedges, downtoners, emphatics, and attitudinal expressions (such as private verbs). Radio interview had the highest proportion of these features while news scripts, perhaps the most 'detached' register in the corpus, consistently had amongst the lowest.

The other dimensions that Biber (1995) lists as 'universal' are as follows: an oral/literate dimension that is specific to each language; one marking narrative registers; one involving the marking of persuasion or argumentation. Without having done a multidimensional register study on Scottish Gaelic, it is impossible to comment on any dimensions that are specific to the language. However, it is clearly the case that there is a distinction in this corpus between narrative and non-narrative discourse, as covered in finding 3 at the beginning of the chapter. In terms of the last 'universal' dimension mentioned by Biber, there were no traces of a persuasive or argumentative dimension in this data set. This is probably due to the absence of registers here that are marked by such a dimension, such as editorial prose, televised debate, or sermons. In future studies, perhaps the inclusion of such registers would lead to the detection of such a dimension in Scottish Gaelic. However, for any such study – and, for that matter, any attempt to assemble a representative corpus of written and spoken Scottish Gaelic – it is important to ask whether or not such registers are an important part of this particular language's register variation.

To conclude this section, many of the findings of past work on register variation were replicated in the present work. In particular, Biber's (1988) observation that the linguistic differences of speech and writing ought not to be discussed in terms of a dichotomy, but as a multidimensional range of variation, is borne out by this study. The few cases of clear opposition between the modes in terms of linguistic features was reduced mainly to production constraints, which differ in 'typical' speech and writing, but not categorically so. This point was raised in §9.2 regarding the differences and similarities of traditional narrative, which is a rehearsed, spoken register, and news scripts, a written register with constrained production time. Many of Ochs' (1979) findings for unplanned language were duplicated here. The main ones were: the lack of explicit marking of semantic relations between referents and following predications; the presence of asyndetic juxtaposition; the lower number of relative clauses; and the lower use of lexical strategies for encoding meaning. However, one major discrepancy between English and Gaelic was the lack of

evidence for the 'historical present' in spoken narrative. Chafe's (1982) dimensions – integration *vs.* fragmentation and detachment *vs.* involvement – were found to be relevant for Scottish Gaelic, particularly between the most opposed registers, conversation and formal prose. Finally, some of the dimensions of variation located in the multidimensional register studies of Biber and others were clearly substantiated for this corpus of Scottish Gaelic. In particular, those dimensions mirroring Chafe's were relevant (e.g. informational *vs.* involved production and ones related to production constraints), as were differences between narrative and non-narrative discourse, and, to some extent, situated *vs.* explicit reference. A future multidimensional study of Scottish Gaelic register variation may detect further dimensions, but it is likely to find the aforementioned one, based upon the present results.

9.4 Concluding Remarks on Spontaneous Spoken Gaelic

In the present study, four spoken registers were included: conversation; radio interview; sports reportage; and traditional narrative. Of these, conversation is the most contextually typical of spontaneous spoken language as described by Miller and Weinert (1998) (see §2.1.3). Traditional narrative, by definition, has been passed down and fine-tuned through many re-tellings and cannot be considered impromptu in the way that the other spoken registers can. As we have seen, sports reportage, although produced on-line, deviates substantially from the other spoken registers both contextually and linguistically; severe production constraints and less discourse freedom are two pervading differences. Radio interview is contextually distinct from conversation by the presence of an interviewer and an audience, as well as the fact that most participants are unable to see each other. These differences may contribute to its greater occurrence level of relative clauses, coordinate clauses, and fronted adverbials and its lower proportion of *'s* clefts and clausal negation. Moreover, a greater self-consciousness arising from its public nature is probably responsible for the lower level of English borrowing evinced than in conversation. In light of these differences and certainly many more than this particular study was capable of detecting, it is important to be aware that media language is not an adequate substitute for genuine private conversation. However, there are a number of linguistic properties shared by the two registers:

- high levels of demonstrative clefts
- presence of detachment constructions
- smaller clauses, as measured by the higher levels for main/matrix clauses
- less complex noun phrase structures
- fewer attributive adjectives (shared by the spoken registers in general)
- lower proportion of nouns
- less conservative grammar
- greater involvement, as evinced by the higher level of interrogative clauses
- lower number of prepositions, indicating a lower informational orientation
- a greater proportion of unassimilated English borrowing than other registers
- fewer long words

Many of these differences relate to those found by Miller and Weinert (1998) in their survey of the spontaneous speech of a number of different languages. In particular, both studies have detected the following properties:

1. Information in spontaneous spoken language is staged, built up over multiple clauses and phrases in 'bite-sized chunks'
2. Syntax tends to be more fragmented, seen in the greater frequency of asyndetic clausal juxtaposition, detachment (or NP-Clause, etc.) constructions, and small main clauses
3. In light of these properties and other facts about spoken language (see Chapter II), the sentence is not an appropriate structure of analysis for register analysis or studies of spontaneous speech

4. Subordinate constructions are generally less frequent in spontaneous speech than in writing; in the current study, this was the case with relative clauses, and to some extent complement clauses, and concessive clauses, but not with temporal or cause and reason clauses
5. Noun phrases tend to be simpler in spontaneous speech, with fewer modifiers such as attributive adjectives and prepositions
6. The structure of modified NPs in spontaneous speech tends to be simpler than in writing
7. Complex noun phrases tend to be avoided in subject position with transitive clauses but are sometimes found in extracausal positions (Miller and Weinert 1998: 141)
8. Lexical variation is narrower in spontaneous speech than in writing

In sum, there was an overwhelming amount of similarity between Miller and Weinert's (1998) results and those of the current study. This bolsters the impression that there are certain features of spontaneous speech which are universal in human language. However, one interesting difference between the studies concerns NP modification. It was found that N-adjective-adjective combinations (ex. *cù mòr dubh* 'a big, black dog') were more frequent than combinations of N-adjective + other type of modifier. In fact, they were more frequent than any other complex noun type in the spontaneous spoken registers. This diverges from the data of Miller and Weinert (1998) and further research would be useful on the subject.

One area which Miller and Weinert examine comprehensively is the tendency for spoken language to evince constructions missing from written language, and vice versa. Beyond the use of juxtaposition, and, to a certain extent, left-detachment, the current study did not find evidence for such clear-cut differences in Scottish Gaelic. However, as this investigation was primarily concerned with linguistic features having occurrence levels above 1 per 1000 words, it did not attempt to catalogue such constructions, which are generally less frequent and more difficult to attend to. Moreover, it is not as clear with Scottish Gaelic where one should begin to look for such differences.

There is a major gap between the level of prescriptivism inculcated in English schools and that present in Gaelic education, which is newer and has fewer resources and literary models for representing the principles of 'educated usage'. Most Gaelic speaking adults, in fact, received almost all of their education in English, and rarely or never write the language. However, because writing generally permits a greater range of syntax and lexis than speech, it seems likely that even inexperienced Gaelic writers will produce linguistic forms reflecting this difference. Some of these will be analogues of English style, others may hark back to older forms of the language, while others will simply be a product of low production constraints. The current study provides a wide net for detecting differences between spoken and written Gaelic. There remains a need for angling of a deeper nature, however, which would uncover those less frequent register 'markers' which immediately declare "this is formal prose" or "this is colloquial conversation". This is one suggestion for future work. Others will follow in the next, and final section.

9.5 Implications and Suggestions for Future Study

Previous investigations on Scottish Gaelic have mentioned ways in which certain registers seem to differ from one another, but these have been based upon largely introspective or anecdotal evidence. As a result, it has remained unclear whether Gaelic, a dying, minority language, evinces the kind of contextually-driven morphosyntactic variation found by researchers on other languages, or suffers from the attenuated or even nonexistent stylistic variation predicted by Dressler (1988). The results of this study provide, for the first time, hard and empirically garnered evidence that Gaelic, despite the contraction of its speakers and geographical base, still displays a range of register variation comparable to that found in English and other larger languages.

Employing features that have been absent from or rarely used in past register research—such as cosubordination, nominal case grammar, lexical borrowing from a higher status language, detachment, and different kinds of highlighting expressions – this study advances current knowledge about the ways in which linguistic form correlates with context. In a general sense, as so many of the findings are similar to those obtained by other researchers working on an array of different languages, they bolster the impression that there are certain 'universal' parameters of register variation. This is especially notable as few of these studies have been on minority languages, and far fewer on ones that are threatened to the point that Gaelic is.

For the Gaelic linguist or grammarian, the investigation provides information on usage and distribution which is almost absent from available descriptions of the language. Furthermore, as the field of Gaelic syntax has been almost entirely concerned with written Gaelic, there is the potential that the results will foster an increase in awareness of structures that are more typical of speech. The study offers a helpful reference point for scholars planning additional research on Gaelic register development and variation. Both studies interested in charting the development of a single register (ex. Lamb 1999) and in more global differences, as in a multidimensional approach, will be aided by the results. Particularly useful will be the tag set developed here, which could be either expanded or compressed as needed. Along with *LinguaStat*, it provides the building blocks of an automatic part of speech tagger and concordance programme, which are two essential tools of corpus linguistics. Finally, far from being the last word on the linguistic features surveyed, the study highlights the need for additional work on the discourse functions, not to mention other syntactic variants (e.g. Gaelic clefts), of a wide range of constructions.

There is an almost limitless range of possibilities for further research on the grammar and discourse of colloquial, spoken Scottish Gaelic. As the language rapidly declines, a rapid response to research in these areas becomes more and more crucial. It is ironic that the type of Gaelic most heard by native speakers themselves as they grew up and most used in the community is the one that linguists and language learners have the least access to and, consequently, understand the least. Currently, the only recordings known to this author of natural, conversational Gaelic are the ones that were completed for this study. There is a dire need for archiving this, in many ways, most important variety of Scottish Gaelic before it is irretrievably lost or unstable. I have articulated one proposal for such an attempt in the past (Lamb 2005), the development of a large corpus of spontaneous, spoken Gaelic. Without a resource of this nature, there is no guarantee that research on natural, colloquial Scottish Gaelic will not, in time, be as extinct as the language itself seems destined to be.

Independent of such an effort, or in coordination with it, one could assemble a representative corpus of the language, including a far wider range of spoken and written registers than was possible here. If these resources were available, they would open up entirely new routes for Scottish Gaelic studies. Biber's work has demonstrated the powerful applicability of corpora to register variation (1988 *inter alia*) and grammatical description (*et al.* 1999). Corpora have been put to good use in machine assisted lexicography, such as in the development of dictionaries and thesauri (see McEnery and Wilson 1996 for a review). With the creation of an automated part of speech tagger for Scottish Gaelic, these possibilities would be greatly facilitated, as would others such as syntactic parsers, speech recognition, and, with a parallel corpus of English language, automatic translation programmes. Such developments would raise the profile of Scottish Gaelic research and offer scholars a wide range of new tools for examining and working with the language. But, perhaps most importantly, they would have broad and lasting benefits for Scottish Gaelic teachers, students, and users in general. In contrast to work of a purely theoretical nature, developments such as these have the potential to benefit a large sector of the Gaelic writing, speaking, and learning public.

Appendix 1:
A Descriptive Grammar of Scottish Gaelic

1 Phonology: A Brief Summary

In contrast to the many phonetic, mainly dialectal, studies that have been undertaken on ScG, there are very few full phonological analyses in the literature (but see Oftedal 1956; Ternes 1973; Dorian 1978; Ó Murchú 1989). However, as Ternes (1994) points out, there are several phonetic studies which, due to their consistent approach, are capable of being phonologically reinterpreted (i.e. Dieckhoff 1932; Borgstrøm 1937; 1940; 1941; Mac Gill-Fhinnein 1966).

1.1 Vowels

The vowel system has three classes – front-unrounded, back-rounded, back-unrounded – and four grades of opening (Ternes 1994):

Simple Vowels

i u u
e ʏ o
ε ɔ
a
[+/- long]

Nasal Vowels

ĩ ũĩ: ũ
õ
ẽ
ã
[+/- long except ʏ]

Vowel length is a major feature in the language, with minimal pairs such as a *car* /kar/ 'a turn or twist' and *càr* /ka:r/ 'car' readily found. In a broad sense (see Ternes 1973 for more information), vowels undergo nasalisation when in contact with nasal consonants (i.e. *m*, *mh*, *n*, *ng*). It seems that nasalisation affects a word *in toto* rather than just the vowel segment (Ternes 1994). Ternes says the following:

The vowels enumerated above function as the *centres of nasalisation*. Every word with distinctive nasalisation contains one such centre, always located in the stressed syllable. From this centre nasalisation usually spreads over several segments to the left and right, but not beyond word boundaries. The precise extension of these nasalized stretches is strictly rule-governed. (p. 102)

An example of this 'spreading' can be seen in a word such as *màthair* 'mother' /mã:har/ [mã:hã]. True minimal pairs are rare, but do occur: *tàmh* /tã:v/ 'rest' and *tàbh* /ta:v/ 'fishing net'.

The Gaelic vowels are assigned to two groups: the so-called 'broad' or back vowels (*a*, *o*, *u*) and the 'slender' or front vowels (*e*, *i*). In the orthography⁴⁶, this distinction is mapped onto the consonants (and many consonant clusters) each of which, save the labials⁴⁷, has a 'broad' (non-palatal) or a 'slender' (palatal) form. For instance, the phonological dental in the word 'cats' is palatal – *cait* /ket^b/ [ke^btʃ] – while it is non-palatal in the singular form of the word: *cat* /kat^b/ [ka^btʃ]. One might be lead, based upon such an example, to claim palatalisation in consonants is a case of assimilation to the vocalic environment, with front vowels assigning palatal allophones and back vowels, non-palatal allophones. However, as is represented in the following two examples, this does not seem to be entirely accurate: cf. *ùr* NOM 'new' [u:r] vs. *ùir* GEN 'new' [u:ɪ]. Although the terminal vowel is phonetically the same in the two word forms, the first realises an unpalatal /r/ while second realises a palatal one.

⁴⁶ There is an orthographic convention that a consonant of one type must always be flanked by vowels of the same type, known as the *slender to slender, broad to broad* rule.

⁴⁷ But see Macaulay (1966).

1.2 Consonants

The phonological consonant inventory varies from dialect to dialect, but the following are widely attested (palatalisation indicated by ‘’):

Table 122: Consonant inventory

	<i>Labial</i>	<i>Dental</i>	<i>Dental’</i>	<i>Velar’</i>	<i>Velar</i>	<i>Glottal</i>
Plosives						
voiced	b	d	d’	g’	g	
voiceless	p	t	t’	k’	k	
Fricatives						
voiced	v					
voiceless	f	s	ʃ	ç	x	h
Nasals	m	n	n’		N	
Vibrants		r	r’		R	
Laterals		l	l’		L	

Most of the Hebridean dialects (e.g. North Uist) present a 3-way contrast in the sonorants but the Gaelic of other areas, particular in the periphery, may range from a one-way contrast (as in the *r* phonemes of East Sutherland Gaelic: see Dorian 1978) up to a 5-way contrast, as with the laterals and nasals of the Islay dialect (Hamp 1970, cited in Ternes 1994). In general, the number of *r* phonemes tends to be fewer than the others, with most dialects evincing between 3 and 2, usually preserving a contrast between palatalised and non-palatalised forms (e.g. the Gaelic of Applecross, Rosshire: see Ternes 1973).

1.3 Diphthongs

Due to the interaction of the features of short *vs.* long vowel length and oral *vs.* nasal articulation, there is a considerable proliferation of distinct diphthongs and triphthongs in Scottish Gaelic. Dialect studies report, on average, between 20–80 diphthongs (Ternes 1994). MacPherson (1940) gives the following base forms for the North Uist dialect:

- 1) ‘Rising’ ai, ei, oi, ui, ʔi, au
- 2) ‘Falling’ iə, ia, iu, eə, eo, ua, uə, uʔ

Common triphthongs include -coi-, -iui-, and -uai- as in the words *eòin* ‘birds’, *ciùil* GEN ‘music’, and *luaithe* ‘faster’.

1.4 Word Stress

Very little has been published on ScG prosody and suprasegmentals. Sound beginnings have been made, however, by Macaulay (1979) and Ó Murchú (1989) and the interested reader should consult these accounts.

Stress variation is of little importance to word formation in Scottish Gaelic. Full stress is nearly always on the first syllable (e.g. *tàmailteachadh* /t̪aːmal’t̪s̪əx̪/ ‘shaming’). Deviations from this are almost always either loans – ex. *buntàta* /bunt̪aːt̪aː/ ‘potato’ – or compounds: *cho-dhùin* /x̪oːv̪uːn̪/ ‘concluded’.

1.5 Consonant Mutation

The Celtic languages are famous for the various ways in which their consonants regularly change in certain conditions. Lenition, sometimes referred to as *aspiration* or *initial mutation*, is the most ubiquitous of these changes in Scottish Gaelic. Synchronically, it is part of the language’s morphophonology and used to mark various aspects of the verbal and nominal system including tense, definiteness, possession, and case agreement, for instance:

Tense

(112)	snàmh	thusa	shnàmh	thusa
	swim-IMP	2S-EMPH	swim-PAST	2S-EMPH
	'swim you!'		'you swam!'	

Definiteness

(113)	<i>craobh</i> 'a tree'; <i>a' chraobh</i> 'the tree'
-------	--

Possession

(114)	<i>a cas</i> 'her foot'; <i>a chas</i> 'his foot'
-------	---

Case agreement

(115)	<i>taobh ceàrr</i> 'wrong side'; <i>air an taobh cheàrr</i> 'on the wrong side'
-------	---

Lenition originated as a phonetic phenomenon early on in the history of the Celtic languages, the product of the weakening of intervocalic consonants both word-medially and at word boundaries. Orthographically, lenition is represented by an *h* after so-called 'lenitable' consonants. However, this is misleading as all consonants⁴⁸, as well as some consonant clusters, change articulation when lenition is applied. Table 123 below, based upon the basic IPA symbols used in Ó Dochartaigh (ed. 1994–7: 123) provides an simplified overview of these changes at a phonetic level, with some of the allophones encountered. This is not an exhaustive list, but rather an illustration of some of the phonetic changes.

Table 123: Lenition of Initial Consonants

Unlenited			Lenited		
<i>Spelled</i>	<i>Non-palatal</i>	<i>Palatal</i>	<i>Spelled</i>	<i>Non-palatal</i>	<i>Palatal</i>
b	b	b	bh	v, β	v, β
c	k	c	ch	x	ç
d	d̪	dʒ	dh	ɣ	ʝ
f	f, ɸ	f, ɸ	fh	-, h	-, h
g	g	ʝ	gh	ɣ	ʝ
l	l̪	ʎ	l	l, ɭ	ɭ
m	m	m	mh	ṽ	ṽ
n	n̪	ɲ	n	n, ɳ	ɲ
p	p	p	ph	f, ɸ	f, ɸ
r	ʀ	ɹ, ø	r	r	n/a ⁴⁹
s	s	ʃ	sh	h	h
t	t̪	tʃ	th	h	h

Palatalisation is another common morphophonemic process, whereby a normally unpalatal consonant or cluster becomes palatalised. Palatalisation is a common feature of case and number marking (see A§2.2.1.2) but also appears in adjectival declension

⁴⁸ Except for the sonorants *l*, *n*, and *r* which may or may not depending, largely, on the age and dialect of the speaker.

⁴⁹ In most dialects, there are no nouns or otherwise lenitable words beginning with a palatal *r*.

(A§2.2.4) and the verbal system. The example using *cait* in A§1.1 above, is one instance of it being used to mark number. The following examples show it being used to mark case:

- (116) mo chas dheas
 1S-POSS foot-FEM-N right
 'my right foot'
- (117) air mo chois dheis
 on 1S-POSS foot-FEM-D right-D
 'on my right foot'

In the orthography, palatalisation is signalled by vowel changes, normally the addition of *i* at the end of a 'broad' vowel. However, more extensive changes sometimes occur as seen in (117) and delineated in Table 124 below.

Table 124: Vowel Changes in Palatalisation (based in part on Ó Maolalaigh 1996)

<i>Change</i>	<i>Example</i>	<i>Change</i>	<i>Example</i>
a > oi	Gall, Goill	ia > èi	grian, grèin
a > ui	ball, buill	ia > eòì	nial, neòil
ea > i	ceann, cinn	ia > ì	sian, sìne
ea > ei	eag, eige	ìo > ì	crìoch, crìch
eò > iùì	ceòl, ciùil	iu > i	fliuch, fliche
eu > èi	geur, gèire	o > ui	roc, ruic
eu > eòì	deur, deòir		

1.6 Preaspiration

Preaspiration is uncommon in the world's languages but found in some Scandinavian languages (e.g. Icelandic) and Scottish Gaelic. It is a short breath-like sound occurring generally before post-vocalic voiceless stops, such as /p/, /k/, and /t/. Preaspiration is not a feature of some of the peripheral ScG dialects such as Arran, parts of Kintyre, and E. Perthshire (seen in the transcriptions of Ó Murchú 1989). In other dialects, there is a range of variation which can be organised into four main types (table from Ó Maolalaigh 1997):

Table 125: Preaspiration after back vowels

<i>Type</i>	<i>Phonetically</i>	<i>Phonemically</i>	<i>Illustrative Dialects</i>
A	[^h p], [^h t], [^h k]	/p/, /t/, /k/ ⁵⁰	Lewis; parts of Sutherland
B	[hp], [ht], [hk]	/hp/, /ht/, /hk/	Aultbea in Ross-shire
C	[hp], [ht], [xk]	/hp/, /ht/, /xk/	Outer Hebrides other than Lewis; Isle of Skye; central Highlands
D	[xp], [xt], [xk]	/xp/, /xt/, /xk/	Western Perth; eastern Highlands; parts of Argyll

⁵⁰ In these dialects, preaspiration is phonetic rather than phonemic. Also, it tends to be weaker and more breathy.

2 Morphology

Although Scottish Gaelic syntax, particular its variation, has not been well covered in the literature, its morphology has been documented in a number of grammar books and articles. Calder (1923; see also Nicholson 1936) is the most complete of the grammars, but is conservative, both in terms of its application and the forms of language it evinces, by today's standards.

2.1 Overview of morphological characteristics

2.1.1 *Synthesis and Fusion*

Scottish Gaelic is a largely fusional language as seen, for example, in its large proliferation of pronominal forms (see also A§2.2.2, and A§2.2.6):

- (118) *air* 'on'; *orm* 'on me'; *ormsa* 'on me [emphatic]'
 (119) *ann* 'in'; *annad* 'in you'; *annadsa* 'in you [emphatic]'; *'nad* 'in your'

Its verbal system is less fusional than French or Spanish but approximates that level in some cases:

- (120) *bhithinn*
 be-INDEF2-1s
 'I would be'
 (121) *seallamaid*
 show/look-IMP-2P
 'Let us show/look'
 (122) *rinneadh* *e*
 do-PAST-IMPERS 3S
 'it was done'

The 2nd indefinite (also known as the 'conditional/habitual'), certain passive forms, and the subjunctive/imperative are the only verb forms currently exhibiting synthesis in Modern ScG. However, there is a growing tendency in the language for these forms to be replaced by periphrastic versions (see A§3.5). First person marking also infrequently appears associated with a vestigial 'present' (see the TENSE-ASPECT-MODALITY system in A§2.3.1.2) in some psychological verbs, almost exclusively in high-register writing and religious discourse:

- (123) *chithim* *cluinneam* *saoilim*
 see-PRES-1S hear-PRES-1S think-PRES-1S
 'I see' 'I hear' 'methinks'

These correspond to the more common usages:

- (124) *chì* *mi* *cluinnidh* *mi* *saoilidh* *mi*
 see-INDEF1 1S hear-INDEF1 1S think-INDEF1 1S
 'I see' 'I hear' 'I think'

2.1.2 *Morphological Processes*

There are a number of ways in which ScG stems can be altered: prefixation; suffixation; stem modification; and suppletion. Although infixation was a feature of Old Irish (see Thurneysen 1993), it is not currently evinced in the language. Neither is reduplication a

productive process at the level of root forms, yet adjectives and interjectives (A§3.4.3) are sometimes, as in other languages, repeated in series for intensification: *duine mòr mòr* 'a big, big man'.

2.1.2.1 Prefixation

Prefixation and suffixation occur in three combinations (Macaulay 1992; see also Gillies 1993): prefix + stem; stem + derivational suffix; and prefix + stem + derivational suffix. There are a large number of productive prefixes in Gaelic, but some have become so inextricably associated with certain stems that they can be considered fused (Calder 1923: §100–101, §146). Stress on the initial syllable indicates that this is the case, such as in *còmh'la*, adv. 'together' (*co* 'with' + *làmh* 'hand'). Prefix status is indicated where stress occurs on another syllable or is shared between the first element and another. The following is a list of some of the more common prefixes:

Adjectives

neo- 'un'; *mì-* 'un'; *so-* denoting goodness, easiness; *do-* denoting badness, difficulty; *ioma(dh)-* 'many'

Examples: *neo-fhreagarrach* 'inappropriate' (<*freagarrach* 'appropriate'); *mì-chùramach* 'careless' (<*cùramach* 'careful'); *so-dhèanamh* 'easily done' ; *do-dhèanamh* 'impossible'; (<*dèanadh* VN., 'doing'); *ioma-dhathach* 'multicoloured' (<*dathach* 'coloured')

Nouns

mòr- 'big'; *fìor-* 'true'; *àrd-* 'high'; *mì-* (as above)

Examples: *mòr-chuid* 'majority' (<*cuid* 'part'); *fìor-uisge* 'spring water' (<*uisge* 'water'); *àrd-sgoil* 'high-school' (<*sgoil* 'school'); *mì-rùn* 'disregard' (<*rùn* 'regard')

Verbs

ath- 'again'; *sìor-* 'continually'; *mì-* as above; *co-* 'with'

Examples: *ath-athris* 'repeat' (<*athris* 'relate/narrate'); *sìor-fhàs* 'ever-grow' (<*fàs* 'grow'); *mì-thoilich* 'displease' (<*toilich* 'pleasing'); *co-dhùin/codhùin* 'conclude' (<*dùin* 'close')

2.1.2.2 Suffixes

There are a number of common suffixes used in inflectional operations, usually involving marking of tense, person, number. See A§2.3.2 on verbal morphology. Derivational suffixes include:

Verbal Noun Endings

-ail; *-adh*; *-amh*; *-ail*; *-inn*

Examples: *cumail* 'keeping' (*cùm* 'keep'); *moladh* 'praising' (*mol* 'praise'); *dèanamh* 'doing' (*dèan* 'do'); *leantainn* 'following' (*lean* 'follow'); *saoilsinn* 'thinking' (*saoil* 'think')

Noun Endings

agentive: *-air/-eir* m., *piobair* 'piper'

diminutive: *-ag/-aig* f., *amaideag* 'female fool'

abstractive: *-achd* f., *siorraidheachd* 'eternity'

Adjectival Endings

The most common is *-(e)ach* but other ones include *-da/-ta/-te*, *-ail/-eil*, *-mhor*

Examples: *sunndach* 'happy' (<*sunnd* 'joy' + *-ach*); *dearbhte* 'certain' (<*dearbh* 'proof' + *-te*); *beòthail* 'lively' (<*beò* 'alive' + *-ail*); *atmhor* 'swelling' (*at* 'swell' + *mòr* 'big')

2.1.2.3 Stem Modification

This is generally isolated to: person, gender, and case marking of certain nouns (see A§2.2.1.2); agreement of the same on attributive adjectives; and the comparative and superlative forms of some adjectives. In all these cases, stem modification takes the form of generally predictable vowel change along with palatalisation. These are not simply cases of suppletion and, due to its prevalence, this kind of stem modification could be said to be fairly productive in the language. For instance:

(a) *fear* m., 'man'

Case	Singular	Plural
Nom.	<i>fear</i>	<i>fir</i>
Dat.	<i>fear</i>	<i>fir</i>
Gen.	<i>fir</i>	<i>fhear</i>
Voc.	<i>fhir</i>	<i>fhearaibh</i>

(b) *geal* 'white': comparative and superlative: *gile*

2.1.2.4 Suppletion

Suppletion is characteristic in the paradigms of certain irregular a) verbs (see Appendix 1 for a chart of these), b) nouns, and c) adjectives:

(a) *rach* 'go'

Past-dependent.:	<i>deach</i>
Past-independent.:	<i>chaidh</i>
Indefl-dependent.:	<i>tèid</i>
Indefl-independent.:	<i>thèid</i>
Verbal noun:	<i>dol</i>

(b) *bean*, f. 'woman'

Case	Singular	Plural
Nom.	<i>bean</i>	<i>mnathan</i>
Dat.	<i>mnaoi</i>	<i>mnathan</i>
Gen.	<i>mnà</i>	<i>ban</i>
Voc.	<i>a bhean</i>	<i>a mhnathan</i>

(c) *math* 'good'; comparative and superlative: *feàrr*

2.1.3 Head/Dependent Marking

Gaelic is a dependent-marking language, seen in the way it treats case marking:

- (125) *bùth* *Dhòmhnaill*
 shop Donald-G.
head dependent
 'Donald's shop'
- (126) *le a cheann*
 with 3S-POSS head-D
head dependent
 'with his head'

As is clearly seen, in ScG, it is not the heads of these dependent relations which exhibit morphological marking, but rather the dependent element. This is signified by lenition and palatalisation in the first example and lenition only, in the second.

2.2 Nominal Morphology

The Scottish Gaelic nominal system declines for case, definiteness, number, gender, and person as well as displaying morphosyntactic contrasts for alienability and count. Lenition and palatalisation take a heavy loading for marking case, gender and definiteness. A levelling of morphological contrasts is occurring in the language due both to internal trends towards simplification and the pervasive influence of English. The effects can appear erratic at times. Use of conservative morphology correlates positively with written and formal registers as well as speaker/writer age. See Gillies (1980), and Macaulay (1982b, 1986) for more information.

2.2.1 Nouns

The nominal grammatical types evinced in ScG are as follows: common nouns; proper nouns; pronouns (including possessive pronouns and prepositional pronouns); anaphoric nouns; mass *vs.* count nouns; alienable *vs.* inalienable nouns; and verbal nouns. A set of derivational suffixes provides the means for coding diminution (e.g. *-an*, *-ag*), abstraction (*-ad*, *-achd*), and indefiniteness (e.g., *cuideigin* f., 'somebody'; *-eigin* 'or another' < *air choreigin*).

1) Common noun stems can be simple or compound (cf. Macaulay 1992: 207). Compound stems are most often formed by noun-noun combinations, adjective-noun ones⁵¹, and/or by adding prefixes and suffixes:

bùth-obrach (*bùth* 'shop' + *obrach* gen. of *obair* 'work')
mòr-shluagh (*mòr* adj., 'big' + *sluagh* 'people')
ro-shealladh (*ro* prep., 'before' + *sealladh* 'view')
bodachan (*bodach* 'old man' + *-an* diminutive suffix)

The stem, or final suffix if present, indicates case and number. Nouns that are headed by another noun, even in compounds, take the genitive case although this is not always heard in informal speech.

2) Proper nouns in Gaelic take the whole range of case marking and can be adjectivally and appositionally modified. Surnames, while growing in frequency due to English influence, are still not the most common way of referring to a member of the community when speaking in Gaelic. Individuals are more likely to be referred to by nicknames, descriptive epithets (such as occupation) and patronymics, which may have been maintained over generations, rather than a fixed family name (Gillies and Matheson 1994).

(127) Dòmhnall Ruadh Chorùna
 Donald red Coruna-G
 'Red Donald of Coruna'

(128) Òrain Dhòmhnail Ruaidh Chorùna
 Songs Donald-G red-G Coruna-G
 'The songs of Red Donald of Coruna'

3) Pronouns (see A§2.2.2)

⁵¹ There is a closed class of adjectives that only occurs pre-nominally such as *deagh*, *droch*, and *fior* (see A§2.1.4). Here, we are referring to those that *normally* come after the noun. This kind of compounding is usually a marker of more literate, especially poetic, registers.

4) Amongst the class of nouns that can act as anaphors (taking their meaning from surrounding discourse) are: *feadhainn* 'some'; *cuid* 'some'; *fear* 'one'; *tè* 'one'; *sin* 'anything'; *sgath* 'anything'; *dad* 'anything'; *rud* 'thing'; *càil* 'anything'.

5) Nouns that cannot be counted are most often either mass nouns, such as *bainne* 'milk' and *min* 'meal', or collective ones such as *crodh* 'cattle' (Macaulay 1992: 207).

6) An alienable *versus* inalienable opposition obtains in the way that ScG deals with possession (see A§2.2.1.5).

7) Gillies (1993: 204) considers the 'verbal noun' in ScG to be 'in the first instance a noun'. Evidence that can be adduced for this includes the following: a) it can function as a subject or object; b) it can take modifiers; c) it can take determiners; d) it can receive case marking; e) it is assigned for gender. However, it clearly functions in ways that prototypical nouns do not: a) it can receive objects; b) it can link with a closed set of prepositions for the encoding of aspect; c) it can take certain modifiers – particularly adverbs – that prototypical nouns cannot; d) it is clearly not time-stable in the way that most nouns are; e) many cannot form plurals. For more information about the status of the verbal noun see A§2.3 on verbal morphology and Payne (1997: 34–38).

2.2.1.1 Number

Although a vestigial system of dual number marking exists there is now only a obligatory contrast in the nominal system between the singular and plural. Plural number in the nominative is coded most often through palatalisation and suffixation but number marking interacts with case and gender (see A§2.2.1.3 below). The most common plural suffix is *-(e)an*.

Dual number marking occurs with some short feminine nouns and is identically to their dative forms (for examples of the dative see A§2.2.1.4)⁵²:

- (129) *mo dhà làimh*⁵³
 1S-POSS two hands
 'my two hands'

2.2.1.2 Gender/Class

Nouns can be either masculine or feminine. It is sometimes possible to ascertain a noun's gender by examining its suffix morphology (see below) or the way it receives case marking, but assignment is largely arbitrary. For example, there is little correlation between grammatical gender and physical one (ex. *boireannach* m. 'woman'). There are a small number of nouns with defective gender marking which may be either masculine or feminine (depending on dialect), or treated differently depending on case. An example of the latter is *muir* 'sea' which is often masculine when nominative and feminine when genitive (*mara* G.). Some of these nouns have roots in an older neuter gender which is no longer attested in the language.

Calder (1923: 76–77) provides useful information about certain stereotypical endings and classes associated with grammatically masculine and feminine nouns. Nouns ending in *-adh* (verbal noun), *-an/-ean* (diminutive), *-as*, *-ach*, *-aiche*, and *-air* are typically masculine. In addition, words denoting '(t)he names of the elements, seasons of the year, days of the week, metals, colours, grains, vegetables, liquors, and timber' (Calder 1923: 76) are usually masculine. Those nouns ending in *-ag* (diminutive), *-achd/-eachd* (abstract), *-ad/-ead* (abstract), *-e* (abstract), or *-ir* (polysyllables only) are generally feminine. Feminine gender is usually associated with the names of nations, celestial objects, musical

⁵² The article in dual marking can be either /an/ or /na/ depending on the dialect.

⁵³ Cf.

<i>an làmh</i>	<i>na làmhan</i>
ART hand-N	ART hands-N
'the hand'	'the hands'

instruments, afflictions, and corpses as well. In addition to these road markers, it is very often the case that feminine nouns end in palatalised, or 'slender', consonants (e.g. *a' chòir* 'the right') whilst masculine nouns are associated with non-palatal, or 'broad', terminal consonants (e.g. *an cor* 'the condition').

Grammarians have traditionally grouped Gaelic nouns into different classes depending on the relative forms of their nominative and genitive singular. The following classification is based on Konstantopoulous (1998), which largely uses Oftedal's (1956) scheme.

Class I: Masculine Nouns Only

- Nominative singular ends in non-palatal consonant (e.g. *fear* 'man')
- Genitive singular formed by palatalisation (e.g. *fir* 'of a man')
- Plurals formed either by palatalisation (making it homophonous with the G sing.) or suffixation: *-ean* or *-an*
- Nearly all monosyllables form the plural by palatalisation
- Some nouns have a genitive plural form homophonous with the nominative singular

ex. *boireannach* 'woman' m., *boireannaich* G, PL
sagart 'priest' m., *sagairt* G, *sagartan* PL

Class II: Feminine Nouns Only

- Nominative singular ends in non-palatal consonant (e.g. *craobh* 'tree')
- Genitive singular formed through palatalisation and sometimes the suffix *-e* in mono- and disyllabic words (e.g. *craoibhe* 'of a tree')
- The suffix *-e* tends to be elided in contact with other vowels: e.g. *air mullach na craoibh* 'àirde' 'on top of the high tree'
- Plural formed by suffixation: *-(e)an*
- Some nouns have a genitive plural form homophonous with the nominative singular

ex. *làmh* 'hand' f., *laimhe* G, *lámhan* PL
tunnag 'duck' f., *tunnaig(e)* G, *tunnagan* PL

Class IIIa: Masculine and Feminine Nouns

- Nominative singular of masculine nouns may end in any consonant
- Nominative singular of feminine nouns ends in palatal consonant
- Genitive singular marked by *-e* and palatalisation on masculine noun endings which are not already palatal
- Plural formed by suffixation: *-(e)an*

ex. *tuil* 'flood' f., *tuile* G, *tuil(t)ean* PL
brùid 'brute' m., *brùide* G, *brùidean* PL

Class IIIb: Identical to IIIa but genitive singular ending *-a* and depalatalisation occurs with palatal terminal consonants

ex. *loch* 'lake' m., *locha* G, *lochan* PL
fuil 'blood' f., *fala*⁵⁴ G

Class IV: Feminine Nouns Only

- Nominative singular ends in palatal consonant
- Genitive singular formed by depalatalisation, syncope, and the suffix *-ach*
- Nominative plural formed by suffix *-achan/-ichean*, *-nnan* or *-(e)an*

⁵⁴ In some dialects, including Uist, many feminine nouns in the genitive singular are marked by the suffix *-adh*; eg. *ceòl na pìobadh* 'music of the pipes'. The dative of certain feminine nouns may also be marked by *-(a)idh*; eg. *aig a' bhùthaidh* 'at the shop' (see ex. (177)).

- ex. *litir* 'letter' f., *litreach* G, *litrichean*
dùthaich 'country' f., *dùthcha* G, *dùthchannan*
bàthaich 'bye' f., *bàthcha* G, *bàthaichean*

Class V: Indeclinable Nouns

- Only marked, if at all, in the plural (i.e. the genitive singular is identical to the nominative singular)
- Plural is formed irregularly or by suffixing: either *-(e)an* or *-achan/-ichean*
- Loans often fall into this class and usually form the plural by *-ichean*
- Nouns ending in /e/ and using *-ichean* to form the plural drop the /e/ if it occurs between two homo-organic consonants, e.g. *coille* 'forest' > *coilltean* 'forests'

Examples: *duine* 'man' m. (NOM AND GEN), *daoine* PL
lighiche 'healer' m. (NOM AND GEN), *lighichean* PL
compiutair 'computer' m. (NOM AND GEN), *compiutairean* PL

2.2.1.3 Definiteness/Referentiality

Gaelic uses only a definite article; indefiniteness is coded by its absence. The marking of definiteness is conditioned by gender, case and number in a system of some complexity. The article always directly precedes the noun and is proclitic to it. It may cause lenition or nasalisation but, unlike Irish Gaelic, nasalisation is not reflected in the orthography. There are different forms of the article depending on whether the first phoneme of the noun is: a) lenitable; b) a vowel; c) a labial consonant; or d) an 's' cluster. No article is used with the vocative.

Table 126–Table 131 below show the permutations of this system. An ⁴ marks those cases where lenition occurs. As mentioned previously (see A§1.2), dentals are 'protected' from article-imposed lenition due to homorganic blocking. When the article ends in a nasal consonant (as in the masculine nominative singular and all genitive plurals) it is assimilated to /m/ before labial consonants. In addition, some dialects (e.g. Uist) feature the assimilation of voiceless consonants to the voicing inherent to nasal environments. Following nasal consonants, as in certain inflections of the definite article (i.e. *an*, *am*, *nan*, *nam*), voiceless consonants in these dialects often become voiced, especially in informal discourse; e.g. *an cù* 'the dog' [ən ku: > əgu:].⁵⁵ Where a noun begins with /s/ followed by either a vowel or one of the sonorants [l, n, r], the article takes the form of *an t-* in certain cases (e.g. *an t-slat* [ə dlaht]) as seen below:

Table 126: The article: masculine noun, consonant-initial

Case	Singular		Plural	
	Indefinite	Definite	Indefinite	Definite
Nominative	ø	an/am	ø	na
Dative	ø	a' ^l /an t-(s-)	ø	na
Genitive	ø	a' ^l /an t-(s-)	^l	nan/nam

⁵⁵ See Borgström (1940) for more information on nasalisation in the dialects of the Outer Hebrides.

Table 127: The article: masculine noun, vowel-initial

Case	<i>Singular</i>		<i>Plural</i>	
	Indefinite	Definite	Indefinite	Definite
Nominative	∅	an t-	∅	na h-
Dative	∅	an	∅	na h-
Genitive	∅	an	∅	nan

Table 128: The article: masculine noun, /f/-initial

Case	<i>Singular</i>		<i>Plural</i>	
	Indefinite	Definite	Indefinite	Definite
Nominative	∅	am	∅	na
Dative	∅	an ^L	∅	na
Genitive	∅	an ^L	l	nam

Table 129: The article: feminine noun, consonant-initial

Case	<i>Singular</i>		<i>Plural</i>	
	Indefinite	Definite	Indefinite	Definite
Nominative	∅	a' ^L /an t-(s-)	∅	na
Dative	∅	a' ^L /an t-(s-)	∅	na
Genitive	∅	na	l	nan/nam

Table 130: The article: feminine noun, vowel-initial

Case	<i>Singular</i>		<i>Plural</i>	
	Indefinite	Definite	Indefinite	Definite
Nominative	∅	an	∅	na h-
Dative	∅	an	∅	na h-
Genitive	∅	na h-	∅	nan

Table 131: The article: feminine noun, /f/-initial

Case	<i>Singular</i>		<i>Plural</i>	
	Indefinite	Definite	Indefinite	Definite
Nominative	∅	an ^L	∅	na
Dative	∅	an ^L	∅	na
Genitive	∅	na	l	nam

As mentioned above, the language is in the process of morphological simplification. The genitive case in particular is inconsistently marked and the nominative (or dative) system is often used in the place of it. Amongst the effects are the use of *na h-* instead of *nan/nam* for definite plural genitives and the absence of lenition in indefinite genitive plurals.

2.2.1.4 Case

There are 4 cases extant in modern ScG: nominative-accusative⁵⁶; dative; genitive; vocative. Nouns in subject and object position take nominative case. Nouns following most simple pre-positions take dative case. Genitive case marking occurs on nouns

⁵⁶ Called thus because there is no contrast between nominative and accusative case marking.

following a small set of simple prepositions (A§2.2.6.1), most complex prepositions (A§2.2.6.3), verbal nouns, and other nouns with which they are in a dependent relation. Vocative case is used most typically to address others and in asseveration (see A§3.4.3). Some examples of genitive case assignment follow:

Genitive marking after the verbal noun:

- (130) tha mi a' sireadh an leabhair bhigh ghuirm
 be-PRES-IND IS PROG searching-VN ART book-G small-G blue-G
 'I am looking for the small blue book'
 cf. *an leabhar beag gorm* (NOM) 'the little blue book'

Genitive marking after another noun:

- (131) doras na sgoile(adh)
 door-N ART school-G
 'the door of the school'
 cf. *an sgoil* NOM 'the school'

When two or more nouns are governed by a preceding noun, only the last is normally marked as genitive:

- (132) doras oifis ceannard na sgoile(adh)
 door-N office-N head-N ART school-G
 'the office door of the head of the school'

Similarly, when a noun is preceded by a preposition (*air*) and that noun (*cas* NOM; *cois* DAT) governs another noun (*bean* NOM; *mnà* GEN) in the genitive, the former is usually⁵⁷ marked as nominative rather than dative:

- (133) thuit a' chlach air cas mo mhnà
 fall-PAST ART stone on foot IS-POSS wife-G
 'the stone fell on my wife's foot'
 cf. *air a cois* 'on her foot'

Case levelling also occurs when a noun governed in the genitive is the object of a following non-finite verb, with which it forms a small clause (cf. structures in (272) and A§3.2.2):

- (134) tha mi ag iarraidh an litir sin a chur thugad
 is-PRES IS PROG wanting-VN ART letter-N that AGR put-VN to-2S
 'I want to send that letter to you'
 cf. *tha mi ag iarraidh na litreach (GEN) sin* 'I am (at the) wanting (of) that letter'

Appositives are always formed in the nominative case rather than the genitive:

- (135) tha Dòmhnall an gobha ag obair sa cheàrdaich an-dràsta
 is-PRES Donald-N ART smith-N PROG work-VN in-ART forge-D right now
 'Donald the smith is working in the forge right now'

The forms of case marking associated with a particular noun depend largely on its declensional category (see A§2.2.1.2 above). In general, the dative uses the nominative form as its base, the first phoneme(s) of which may be altered following the rules of lenition and affixation as presented above. The endings of feminine nouns in the dative undergo palatalisation where possible (some dialectal variation occurs on this point),

⁵⁷ However, older and very formal texts (e.g. the Bible) preserve dative marking in this situation.

although younger and/or less conservative speakers may not evince this. Most nouns have genitive forms that differ from their nominative forms to greater or lesser extents. Stem modifications – through fairly predictable vowel changes and palatalisation – are frequently all that characterise the genitive forms, but there are a number of nouns which behave irregularly or feature suppletion (see A§2.1.2.4).

Adjectives mark case in a similar fashion to their head noun, through lenition and a certain degree of vowel harmony. Adjectives modifying plural nouns have the suffix /a/ appended to them if their form permits it. See A§2.2.4 for more information about adjectives.

The paradigms presented below are based upon a conservative morphology. Register, proficiency level, and speaker age may be associated with deviations from this standard. Examples in italics offer common reduced forms. Bold text is used to highlight the main features of the system. Where it is said below that certain conditions of case, gender, number, or definiteness indicate lenition, palatalisation, or suffixation, this is to be understood as only being realised in words that have the potential to show these contrasts. For example, words that are vowel initial are unable to evince lenition, words that end in a palatal consonant in the nominative will be unable to evince palatalisation in the genitive⁵⁸, and words already ending in a vowel will not take suffixation to indicate the plural adjectival form.

I. Masculine Noun + Adjective: *fear* 'man'; *mòr* 'big'; *le/leis* 'with'; *taigh* 'house'

Case	Indefinite	Definite
Nominative	<i>fear mòr</i>	<i>am fear mòr</i>
Dative	<i>le fear mòr</i>	<i>leis an fhear mhòr</i>
Genitive	<i>taigh fìr mhòir</i> <i>taigh fear mòr</i>	<i>taigh an fhìr mhòir</i> <i>taigh an fhear mhò(i)r</i>
Vocative	<i>fhìr mhòir</i>	N/A

The base form is the indefinite, nominative *fear mòr* '(a) big man'. In masculine nouns, definiteness in the nominative does not indicate lenition, either in the head noun or the attributive adjective (as it does with female nouns) so we obtain *am fear mòr* (cf. below FEM *a' chaileag mhòr*). The dative form when definite *does* indicate lenition, on both the noun and the adjective: *leis an fhear mhòr*. Much of ScG's nominal morphology, although complex on the surface, is redundant, and parallels obtain between different permutations of number, definiteness and case. An example of this is found in the previous two examples. Mnemonically, one can pair up the morphology of singular masculine nouns that are definite and dative, and singular feminine nouns that are definite and nominative.

The genitive in singular masculine nouns is often characterised by palatalisation (see A§2.2.1.2). Adjectives following masculine nouns in the genitive are lenited and palatalised. When definite, these nouns are also lenited, with an identical form in the vocative, save for the absence of the article.

⁵⁸ Although some of these nouns are depalatalised to show the contrast (see A§2.2.1.2): e.g., *fuil* 'blood' F, NOM; *fala(dh)* GEN.

II. Plural masculine noun + Adjective: *fir* 'men'

Case	Indefinite	Definite
Nominative	<i>fir mhòra</i>	<i>na fir mhòra</i>
Dative	<i>le fir mhòra</i>	<i>leis na fir mhòra</i>
Genitive	<i>taigh fhear mòra</i> <i>taigh fir mhòra</i>	<i>taigh nam fear mòra</i> <i>taigh na fir mhòra</i>
Vocative	<i>fhearaibh mòra</i>	N/A

Type I masculine nouns form their plural by palatalisation and adjectives following these nouns are always lenited. Most adjectives agree with plural nouns by the addition of a vocalic suffix. When indefinite and genitive, all plural nouns are lenited. Type I masculine nouns often revert back to their singular form when genitive. In other masculine nouns, the plural form does not change as a result of case. In the vocative, all nouns are lenited but many Type I nouns additionally evince the suffix *-(a)ibh*, as above.

III. Feminine noun + Adjective: *caileag* 'girl'

Case	Indefinite	Definite
Nominative	<i>caileag mhòr</i>	<i>a' chaileag mhòr</i>
Dative	<i>le caileig mhòir</i> <i>le caileag mhòr</i>	<i>leis a' chaileig mhòir</i> <i>leis a' chaileag mhòr</i>
Genitive	<i>taigh caileig(e) mòir(e)</i> <i>taigh caileag mhò(i)r</i>	<i>taigh na caileig(e) mòir(e)</i> <i>taigh na caileag mhò(i)r</i> <i>taigh a' chaileag mhò(i)r</i>
Vocative	<i>a chaileag mhòr</i>	N/A

Singular feminine nouns in the nominative always indicate lenition, where possible, in attributive adjectives. When definite, singular feminine nouns are also always lenited, except in the genitive. In the dative case, both the noun and adjectives are palatalised. The genitive case, while neutralising lenition, evinces palatalisation. A vocalic suffix is also added to genitive nouns and adjectives if they are 2 syllables or less in length. With those over 3 syllables, these are generally omitted. The vocative is lenited, but not palatalised as in masculine nouns.

IV. Plural feminine noun + Adjective: *caileagan* 'girls'

Case	Indefinite	Definite
Nominative	<i>caileagan mòra</i>	<i>na caileagan mòra</i>
Dative	<i>le caileagan mòra</i>	<i>leis na caileagan mòra</i>
Genitive	<i>taigh chaileagan mòra</i> <i>taigh caileagan mòra</i>	<i>taigh nan caileagan mòra</i> <i>taigh na caileagan mòra</i>
Vocative	<i>a chaileagan mòra</i>	N/A

The marking of feminine plurals and dependent adjectives is identical to the masculine in every respect except one: no feminine plural forms indicate lenition with following adjectives as some Type I masculine nouns do (due to palatalisation, e.g. *fir mhòra* 'big men'; see above).

2.2.1.5 Possession

An alienable/inalienable contrast obtains in the way that Gaelic treats possessions. Presenting a type of iconicity found in many other natural languages, inalienables take a shorter form while alienables take a longer, periphrastic one. This contrast is not so marked in contemporary ScG, but certain regularities are still to be found. In general,

referents which are felt to be 'closer' to the possessor employ the possessive pronouns: e.g. family, pets, one's house, one's clothes, one's wife (but, interestingly, not one's husband in some dialects). The possessive pronoun directly precedes the word it modifies.

Table 132: Possession

Person/Gender	Consonants		Vowels	
	Singular	Plural	Singular	Plural
1	mo ^L	ar	m'	ar n-
2	do ^L	ur	d'	ur n-
3M	a ^L	an/am	a	an
3F	a	an/am	a h-	an

- (136) mo cheann m'obair
 1SG-POSS head 1SG-POSS-work
 'my head' 'my work'

Vagueness can result, especially in orthography, in the 3SG-POSS if the modified noun begins with a non-lenitable consonant or cluster. Older speakers will often preserve a spoken contrast between lenited and unlenited forms of the sonorants (see A§1.5), but with most younger speakers the following would be homophonic: *a làmh* his/her hand.

In cases where the possessor of an inalienable is fully specified, a genitival construction is sometimes used:

- (137) taigh Theàrlaich
 house Charles-G
 'Charles' house'

- (138) còta mòr Màiri⁵⁹
 coat big Mary-G
 'Mary's big coat'

Alienables most frequently are used with a locative construction to express possession: DEF ART + N + *aig* ('at') + POSSESSOR. For instance, Mary's boyfriend would be *an gille aig Màiri*, literally 'the boy at Mary'. If the POSSESSOR is pronominal, one of the forms of the prepositional pronoun derived from *aig* is used (see §2.1.6.1). For instance, 'our television' would be *an telebhisean againn*, while 'my husband' would be *an duine agam*. There is a tendency amongst younger and more progressive speakers to use the periphrastic construction solely.

There is no verb TO HAVE in Gaelic. To express this notion, one uses a locative construction, as above with the verb 'to be'.

- (139) tha càr againn
 be-PRES car at-1P
 'we have a car' [Lit. there is a car at us]
- (140) bha an iuchair agam
 be-PAST ART key at-1S
 'I had the key' [Lit. there was a key at me]

⁵⁹ Note that some dialects evince lenition of feminine proper nouns, so in this case the form obtained would be *còta mòr Mhàiri*.

2.2.2 Pronouns

As was seen in the previous section, pronouns have combined with other elements in ScG to form hybrids: prepositional pronouns and prepositional possessive pronouns. These are part-and-parcel of many idiomatic expressions. Also, like French and German, Gaelic has a limited system of honorifics in the 2nd person. Although there is considerable idiolectal and dialectal variation (see Macaulay 1982a), the general pattern is that *thu* forms are used with familiars and/or contemporaries while *sibh* forms are used with unfamiliar persons and/or elders.

2.2.2.1 Personal Pronouns

Pronominals can be given emphasis by using a number of suffixes. The emphatic suffix is often used for providing contrast or for highlighting purposes, as in 'that's *my* money' (see A§3.4.2.1). Here is a chart of the free pronouns in ScG with their emphatic suffixes in parentheses:

Table 133: Personal PN and emphatic suffix

Person/Gender	Singular	Plural
1	mi(se)	sinn(e)
2	thu(sa)	sibh(se)
3M	e(san)	iad(san)
3F	i(sc)	iad(san)

Depending on the semantic role of a pronominal referent and how it is coded by the morphosyntax of a particular expression, there are cases when one of the following will be used instead of a simple pronoun: a) a possessive pronoun; b) a prepositional pronoun; or c) a prepositional possessive pronoun. The most frequent occurrence of this is with pronominal objects to a verbal noun, as demonstrated in the following examples:

- (141) feumaidh tu a cheannsachadh
 must-INDEFI 2S 3M-POSS taming-VN
 'you must tame/control him'

Here, the modal verb *feumaidh* 'must', a one-argument predicate, takes a transitive verbal complement *ceannsachadh*. The subject of the complement is omitted through complement argument omission (it is understood as being the same as the subject of *feumaidh*). The object, however, is incorporated in the form of a possessive pronoun. Literally, this clause could be stated as 'you must his taming'. See (272) and A§3.2.2 for more information on these kinds of constructions.

- (142) chaidh an ceannsachadh agad
 go-PAST 3PL taming-VN at-2SG
 'they were tamed by you' [Lit. 'went their taming at you']

Similar to example (142), *ceannsachadh* is the complement of another verb, *chaidh* 'went'. The verb 'to go', when taking a verbal argument in this way, has the connotation of passivisation. To express the ACTOR in these constructions, a locative (*aig*) is used or, increasingly in modern usage, the preposition *le* 'with/by'.

- (143) tha iad gad iarraidh
 be-PRES 3PL PROG-2POSS wanting
 'They are wanting you' [Lit. 'they are at your wanting']
 (cf. **Tha iad ag iarraidh thu*)

In this clause, the subject is expressed as a personal pronoun but the object of the verbal noun must be expressed as a possessive prepositional pronoun. The progressive particle in ScG is *ag* 'at', which has its roots in the cognate preposition *aig*. The form *gad* is a fusion of this particle and the 2nd person possessive pronoun *do* 'your'.

2.2.2.2 Demonstratives

ScG demonstratives encode three degrees of proximity or specificity:

Table 134: Demonstratives

Degree	English	PN/ADJ	ADV _{movement}	ADV _{location}
1	this/here	<i>seo</i>	<i>an-seo</i>	<i>an-sheo / ann a' sheo</i>
2	that/there	<i>sin</i>	<i>an-sin</i>	<i>an-shin / ann a' shin</i>
3	yon/yonder	<i>siud/ud</i>	<i>an-siud</i>	<i>an-shiud / ann a' shiud</i>

Demonstrative pronouns have a similar distribution to common nouns; they never combine with prepositions the way that pronouns do. Demonstrative adjectives, like other adjectives, follow the noun they modify (e.g. *an duine ud* 'yon man'). The demonstrative pronoun is the same as the demonstrative adjective in the first two degrees of proximity, but deviates in the third: *chì mi an duine ud* 'I see yon man' *vs.* *chì mi siud* 'I see that'. As can be seen in Table 134 above and in the following example, demonstrative adverbs in some dialects mark a contrast between location and movement: the location forms are lenited. Consider the following:

- (144) chaidh sinn an-sin bha sinn an-shin
 go-PAST 1PL there be-PAST 1PL there
 'we went there' 'we were there'

Although degrees 2 and 3 are sometimes used interchangeably, when a difference is maintained between the two, it seems to be mostly down to visibility or accessibility. Whereas *an cnoc ud* 'yon hill' would refer to a hill that is either obscured from view or just visible, *an cnoc sin* would refer to one that is much more discernible and able to be specifically pointed out. If one were to say *an latha ud* 'yon day', the day would probably not be specifically fixed in the mind of the addresser whereas *an latha sin* 'that day' would imply that it is accessible.

A very natural occurrence in Gaelic is the pairing of tense and demonstratives, which sometimes has the function of introducing referents that are to play a major role in following discourse (see A§3.4.2):

- (145) bha sinn ann air an latha a bha seo
 be-PAST 1PL EXIST on ART day-D REL was here
 'we were there on this day' [Lit. 'we were in it on the day that was here']

2.2.2.3 Reflexives and Reciprocals

Reflexives

There are few lexical reflexives in the language akin to the sort 'I shaved' or 'I washed' where the AGENT and PATIENT roles refer to the same entity. Most reflexive operations involve the use of *fhèin/fhìn*, a fairly prototypical analytic reflexive equivalent to English 'self':

- (146) chunnaic mi mi fhèin⁶⁰
 see-PAST 1SG 1SG REFL
 'I saw myself'

Fhèin/fhìn can also be used as a prefix (e.g. *fhèin-eachdraidh* 'self' + 'history' > 'biography') and to intensify or provide contrast:

- (147) tha mi fhèin sgìth
 be-PRES 1SG REFL tired
 'I myself am tired' (although you are not) or 'I am tired also' (although it seemed as if I wasn't)
- (148) dhèanadh amadan fhèin sin (from Macaulay 1992: 197)
 do-INDEF2 fool REFL that
 'even a fool could do that'

Finally, it can be used to show co-reference between the possessor of an NP and a verbal argument:

- (149) Chuala mi mo bhràthair fhèin
 near-PAST 1SG 1SG-POSS brother REFL
 'I heard my own brother'

There is a morphological reflexive construction employing the verbal noun where the SUBJECT and the SUFFERER – which is encoded on the pre-VN aspect marker – are the same:

- (150) tha sinn gar losgadh
 be-PRES 1PL PROG-1PL burning-VN
 'we are being burnt'

Macaulay calls this Scottish Gaelic's only 'genuine passive' (see A§2.3.2.2).

Reciprocals

There are lexical reciprocals such as 'they kissed' (*phòg iad*) where reciprocity is semantically encoded, but, like lexical reflexives, they seem to be more rare than in English. Often the strategy is for the analytic reciprocal *a chèile* to be used, such as in:

- (151) marbhaidh sibh a chèile
 kill-INDEF1 2PL RECIP
 'you will kill each other'
- (152) bhruidhinn iad ri chèile
 speaker-PAST 3PL to RECIP
 'they spoke to one another'

As seen in the second example, prepositions can be inserted directly before the reciprocal, eliding the *a*.

2.2.2.4 Interrogative Pronouns

There is a small set of question words for requesting information more elaborate than simple affirmation or disaffirmation. These occur utterance initial and directly before the verb. One (*càite* 'where') takes the 'dependent' form of the verb while the others take the

⁶⁰ Most dialects use *fhìn* with 1st person pronouns, eg. *chunnaic mi mi fhìn* or *chunna sinn sinn fhìn*.

'independent' form (see A§2.3.1). In those taking independent forms, the relative pronoun *a* is present, but may be elided in speech before the verb if there is vowel contact. This happens throughout the spoken language; whenever two vowels are in contact, the first usually elides the second.

Unlike English, most of the interrogative forms are different from equivalent relative forms (see Ó Maolalaigh 1996 for more information):

Question Word	Gloss	Relative Form
<i>càite</i>	where (<i>cia</i> 'what' + <i>àite</i> 'place')	<i>far</i>
<i>cò</i>	who	<i>cò</i>
<i>dè, gu dè</i>	what (<i>ciod</i> 'what' + <i>è</i> 'it')	<i>na</i>
<i>carson</i>	why (<i>cia</i> 'what' + <i>a(i)r son</i> 'for')	<i>airson</i> ('s), <i>et al.</i>
<i>cuine</i>	when (<i>cia</i> 'what' + <i>ùine</i> 'time')	<i>nuair</i>
<i>ciamar</i>	how (<i>cia</i> 'what' + <i>mar</i> 'as')	<i>mar</i>

Some examples of each follow:

	Question Form	Relative Form
(153)	<i>càite</i> an d'rinn thu sin? 'where did you do that?'	chunnaic mi <i>far</i> an d'rinn thu sin 'I saw where you did that'
(154)	<i>cò</i> (a) rinn sin? 'who did that?'	chunnaic mi <i>cò</i> (a) rinn sin 'I saw who did that'
(155)	<i>dè</i> (a) rinn thu? 'what did you do?'	is toil leam <i>na</i> (a) rinn thu 'I like what you did'
(156)	<i>carson</i> a rinn thu sin? 'why did you do that?'	rinn thu sin <i>airson</i> 's gum faiceamaid e 'you did that so that we would see it'
(157)	<i>cuine</i> a rinn thu sin? 'when did you do that?'	rinn thu sin <i>nuair</i> a bha sinne air falbh 'you did that when we [EMPH] were away'
(158)	<i>ciamar</i> a rinn thu sin? 'how did you do that?'	chì mi <i>mar</i> a nì thu e 'I see how you do it'

2.2.2.5 Indefinite Pronouns

There are several ways to encode non-specificity or indefiniteness. With nouns, the usual way is to use the modifier *air choreigin*:

- (159) *cù* *air choreigin*
dog 'or another'
 'some dog or another'

There is also a small number of words that can be translated as 'something' or 'nothing (at all)': *càil*, *sgàth*, *sion*, *dad*. The phrase *sam bith* (~ 'in the world'), translatable as 'at all', can modify NPs to convey indefiniteness: e.g. *rud sam bith* 'anything at all'. The generic forms of the masculine and feminine nouns, *feàr* and *tè* respectively, can be used anaphorically to indicate 'one'. Other constructions are as follows, generally splitting into those that prepose the pronominal element (*air bith*) and those that postpose it (*ge be air bith* etc.):

cò/ càite/ cuine + air bith

'who(m)ever' / 'wherever' / 'whenever'

ge (be air) bith + cò/ càite/ cuine/ ciamar

'whoever' / 'wherever' / 'whenever' / 'however'

ge brith + (as previous example)...

as bith cò + etc...

ge b' e cò + etc.

2.2.3 Numerals

The area of ScG number is very interesting but available space allows only a cursory treatment. Gaelic is one of the few reported languages with a vigesimal system (base 20). It has native terms for numbers 1 through 999,999 (Old Irish *mìle* '1000' probably being an early borrowing from Latin).

2.2.3.1 Cardinal Numerals

There are discrete words for numbers 1 through 10, and for 20, 100, and 1000; other numbers are built through compounding. For the purposes of reciting numbers, as in counting or giving telephone numbers (only really done on the radio, where presenters try to adopt an all-Gaelic style), a particle is used: *a h-* before vowels and *a* before all others. Also, in this case the number two is lenited, becoming *a dhà*. The following chart provides the cardinal numbers from 1 to 20 with an accompanying noun:

1	aon(a) ^(L) (chù)	<i>one dog</i>	11	aon (chù) deug
2	dà ^(L) (chù)	<i>two dog</i>	12	dà (chù) dheug
3	trì ^(L) (coin)	<i>three dogs</i>	13	trì (coin) deug
4	ceithir ^(L) (coin) ...		14	ceithir (coin) deug
5	còig ^(L) (coin)		15	còig (coin) deug
6	sia (coin)		16	sia (coin) deug
7	seachd (coin)		17	seachd (coin) deug
8	ochd (coin)		18	ochd (coin) deug
9	naoi (coin)		19	naoi (coin) deug
10	deich (coin)		20	fichead (cù) <i>twenty dog</i>

As can be seen, some numbers cause lenition. Some also take the singular forms of a noun, notably 1, 2, and 20 and multiples of 20 and 100. When constructing numbers between 20 and 40, there are two strategies: a) using the preposition *air* 'on' or *thar* 'past'; b) adding *agus* or 's 'and' and then the secondary number. Above 40, the latter is used:

21	aon(a) (chù) air/thar fhichead	fichead (cù) agus a h-aon
22	dà (chù) air/thar fhichead	fichead (cù) agus a dhà
23	trì (coin) air/thar fhichead	fichead (cù) agus a trì
39	naoi (coin) deug air/thar fhichead	fichead (cù) agus a naoi deug
40	dà fhichead (cù)	<i>identical at this point</i>
50	leth-cheud cù	...
51	leth-cheud cù 's a h-aon	
99	ceithir fichead (cù) 's a naoi deug	
100	ceud cù	
200	dà cheud cù	
1000	mìle cù	
1000k	mìlleann cù	

This system of counting is felt by native Gaels to be cumbersome for larger numbers although the fact that all adult Gaelic speakers received their mathematical education through English has much to do with this.⁶¹ Six thousand six hundred and ninety three (6693) is rendered as *sia mìle sia ceud ceithir fichead 's a trì deug* 'six thousand six hundred four score and thirteen'. This is only slightly longer than the English version, but native Gaelic speakers generally revert to English numbers for those greater than twenty that fall between the scores. For instance, they may say '*naoi fichead caora*' for 180 sheep, but if they needed to be more specific, say, 187 of them, they would probably say 'one hundred eighty seven *dhiubh*'. Dates and phone numbers are usually given in English unless a speaker is trying to avoid code switching. See Macaulay (1982a) for more information on the sociolinguistic aspects of Gaelic numerals.

2.2.3.2 Ordinal Numerals

Gaelic has native terms for the ordinal numbers although their use is quite restricted. Ordinal numbers above 10 are often given in English. Here are the first 10, including 20 and 21:

1 st	a' chiad (chù)
2 nd	an dàrna/dara (cù)
3 rd	an treas(amh)/tritheadh (cù)
4 th	an ceathramh (cù)
5 th	an còigeamh (cù)
6 th	an siathamh (cù)
7 th	an seachdamh (cù)
8 th	an t-ochdamh (cù)
9 th	an naoidheamh (cù)
10 th	an deicheamh (cù)
20 th	am ficheadamh (cù)
21 st	an t-aona (chù) fichead

2.2.3.3 Numerical Pronouns

These are used for referring to an individual or to groups and are declined for gender:

1. aonan/aonar m.
2. dithis(t) f.
3. trìuir f.
4. ceathrar m.
5. còignear m.
6. sianar m.
7. seachdnar/seachdar m.
8. ochdnar/ochdar m.
9. naoinear m.
10. deichnear/deineir m.

When used to modify a noun, those above *aonan* govern the genitive case: e.g. *dithis fhear* two men.

2.2.4 Adjectives

Adjectives in Gaelic cover a wide range of properties ranging from age and value judgements to physicality and shape. Adjectives can be both attributive and predicative. Postnominal attributive adjectives, but not predicative ones, agree with their nominal heads in number, case, and gender. Their distribution indicates that they are

⁶¹ A decimal system has been introduced in the schools, based upon the Irish system.

morphosyntactically distinct from verbs and nouns but there are some which can be nominalised such as *na mairbh* 'the dead' (adj. *marbh* 'dead') or *na big* 'the meek' (adj. *beag* 'small').

2.2.4.1 Prenominal Attributive Adjectives

As mentioned there is a set of adjectives that usually or always occur before the noun. These are all mono- or di-syllabic and communicate basic semantic properties. Although there are other adjectives that sometimes occur pre-nominally (particularly colours), their use in this way is confined largely to poetic language.⁶² Prenominal adjectives always lenite a following noun except where there is homorganic blocking:

<i>ath</i>	next	<i>an ath dhuine</i>	'the next man'
<i>corra</i>	odd	<i>an corra nighean</i>	'the odd girl' (as in 'one here and there')
<i>deagh</i>	good	<i>deagh latha</i>	'a good day'
<i>dearbh</i>	same	<i>an dearbh bheachd</i>	'the same opinion'
<i>droch</i>	bad	<i>droch naidheachd</i>	'bad news'
<i>fìor</i>	true	<i>fìor charaid</i>	'a good friend'
<i>sàr</i>	excellent	<i>sàr obair</i>	'excellent work'
<i>seann</i>	old	<i>seann taigh</i>	'old house'

When definite, these occur after the article and, as they seem to nearly compound with the following noun, they are subject to the initial mutation and affixation that occur with nouns in general. For instance, 'the end of the next month' is usually rendered as *deireadh na h-ath mhìos* (*mìos* f. 'month') while 'the end of the next day' would be rendered *deireadh an ath latha* (*latha* m. 'day'). However, their presence sometimes neutralises case and gender marking, as in *an ath dhuine* (*duine* m. 'man'). Here, one would have expected *an t-ath dhuine* but this is rare.

2.2.4.2 Postnominal Attributive and Predicative Adjectives

These occur directly after the head of the noun phrase. Macaulay (1992: 201) tells us that they form different subclasses and usually occur in the following order:

		size - quality - colour		
(160)	<i>bàta</i>	<i>beag</i>	<i>snog</i>	<i>geal</i>
	boat-MASC	small	nice	white
	'a small nice white boat'			

As is apparent here, all attributive adjectives must agree with the head noun for gender, number, and case. If the noun is in the plural the adjectives change accordingly:

(161)	<i>bàtaichean</i>	<i>beaga</i>	<i>snoga</i>	<i>geala</i>
	boats	small-PL	nice-PL	white-PL
	'nice small white boats' (notice that in English, quality precedes size)			

Predicative adjectives do not agree with their head nouns:

(162)	<i>tha</i>	<i>na</i>	<i>bàtaichean</i>	<i>beag</i>	<i>snog</i>	<i>agus</i>	<i>geal</i>
	be-PRES	ART	boats	small	nice	and	white
	'the boats are small, nice and white'						

See the tables in A§2.2.1.4 for the declension of adjectives with nouns of various case, number, and gender.

⁶² Macaulay (1992, 191) says that the class of pre-nominal adjectives is 'virtually confined' to *deagh*, *droch*, *fìor*, and *seann*.

2.2.4.3 *Quantitative Adjectives*

The words *gach* and *a h-uile* are used to convey the notions of 'every' and 'all' respectively:

- (163) bha a h-uile fear riamh ann
 be-PAST every man ever 'in it'
 'every single man was there'

- (164) gach tè a bh' ann
 every one-F REL be-PAST 'in it'
 'every female that was there'

As can be seen, these act like prepositive adjectives but do not lenite. They require the singular forms of the nouns they modify. *Uile(ag)* can also be used postpositively with a slightly different connotation:

- (165) bha na fir uile(ag) ann
 be-PAST the men all EXIST
 'the men were all there'

2.2.4.4 *Comparison of Adjectives*

Most adjectives have a comparative form which: a) is derived from the base form through stem modification (namely palatalisation and vowel changes) and lenition; b) is identical to the base form; or c) features suppletion. For regularly-formed adjectives, the morphological alterations are similar to those occurring in Type 2 (female) nouns in the genitive singular, i.e. palatalisation and suffixation of /e/ to the final consonant, if present. For instance, the adjective *àrd* 'high' becomes *àirde* 'higher/highest'. However, in polysyllabic adjectives, syncope occurs: *milis* 'sweet' > *milse*; *bòidheach* 'beautiful' > *bòidhche*.

Payne (1997: 88–89) tells us that grammaticalised comparative constructions have three crucial elements: "(1) the known **standard** against which the subject of the clause is compared; (2) the **marker** that signals that the clause is a comparative construction; and (3) the **quality** by which the subject is compared with the standard." Consider the following ScG example (adapted from an Irish example in Payne 1997: 89⁶³):

- (166) tha an cù nas motha na an cat
 COP ART dog MKR bigger-QUAL than ART cat-STD
 'the dog is bigger than the cat' (= 'dogs are bigger than cats')

The marker is *nas* which contains the copula *is* and is therefore sensitive to qualities of tense (see A§2.4.1). This word is followed by the comparative form of the adjective. The quality which is compared here is LARGENESS and *mota* is the comparative form of the irregular adjective *mòr* 'big'. The standard, reflecting trends in the language for predicates to postpose, occurs last in the construction.

The previous example uses the *tha* form of the verb 'to be'. If the copula is used, the form changes slightly:

- (167) is mise as fheàrr na thusa
 COP 1SG-EMPH MKR better-QUAL than 2SG-EMPH
 'It is me who is better than you'

⁶³ Payne's labels for the Irish example are misplaced; they ought to parallel the ones here.

This example would be rare in spontaneous speech today, which would prefer the following: *tha mise nas fheàrr na thusa* (similar to (166)).

The superlative is coded syntactically and is very similar to example (167):

- (168) is mise as fheàrr
 COP 1SG-EMPH MKR-COP-REL better-QUAL
 'I am the best'

With conservative contexts and speakers, the markers tend to agree in tense (being composed of a copular relative) with the main verb:

- (169) bha mise na b' fheàrr na thusa
 be-PAST 1SG-EMPH MKR COP-PAST-REL better than 2SG-EMPH
 'I was better than you'

However, in younger and more progressive speakers particularly – and more often in spontaneous speech than writing – this distinction is not usually upheld:

- (170) bha mise nas fheàrr na thusa

The syntax for these expressions can be very flexible, particularly as displayed in poetry and proverbs. Utterances like the following oral proverb, showing a comparative complement, are characteristic of carefully edited written language (example from Calder 1923: 111):

- (171) Is i a' bhò fhèin as luaithe a mhothaicheas d' a laogh
 COP 3F the cow REFL quickest who notices to her calf
 'the cow herself is the first to notice her (own) calf'

There are two equative constructions depending on whether the standard is nominal or verbal. Both use the word *cho* meaning 'as':

a) Nominal Standards

- (172) tha Màiri cho bradach ri Seumas
 be-PRES Mary as thievish as James
 'Mary is as thievish as James'

b) Verbal Standards

- (173) chan eil Màiri cho bradach agus a tha Seumas
 NEG be-PRES-DEP Mary as thievish and REL be-PRES James
 'Mary isn't as thievish as James is'

2.2.5 Adverbs

The definition of a grammatical category 'adverb' in Gaelic as in other languages is slightly problematic and it can end up as a dustbin denomination. There is no one distinct derivational process associated with adverbs. They either take the form of: non-inflecting words and fixed phrases; or an adjective modified by the particle *gu*, which prefixes *h-* to vowels. They are the most syntactically free elements in the clause and can take various levels of its logical structure in their scope. Probably the easiest way to sort adverbs in Gaelic is by function. The types include adverbs of manner, time, direction/location, and degree.

2.2.5.1 *Manner Adverbs*

This is the largest of group of adjectives as most adjectives can form a manner adverb by the preposing of *gu*, e.g. *trom* 'heavy' > *gu trom* 'heavily'.

<i>gu mòr</i>	greatly
<i>gu h-aineolach</i>	ignorantly
<i>gu coibhneil</i>	kindly
<i>gu fuar</i>	coldly

Like adverbs in general, they have a good deal of syntactic freedom:

- (174) *thuirt i ris gu coibhneil*
 say-PAST she to him ADV kind
 'she said kindly'

cf.: *gu coibhneil thuirt i ris; thuirt i gu coibhneil ris*

All adverbs of this type can be modified using *glè^L* 'very', *ro^L* 'extremely' and other intensifying adjectives: e.g. *gu glè choibhneil* 'very kindly'. There is also a large number of manner adjectives formed by a preposition^(L) + noun. Some examples:

<i>air èiginn</i>	with difficulty
<i>fa sgaoil</i>	in dispersed fashion
<i>an làthair</i>	present
<i>air chall</i>	lost
<i>air seòl</i>	sailing
<i>air iteal</i>	flying

Calder (1923: 309–313) lists many more of these.

2.2.5.2 *Time Adverbs*

These are a well-defined class, dealing with present, past, future, and indefinite tense. Some examples:

PRESENT		PAST	
<i>an ceart-uair</i>	just now	<i>an-dè</i>	yesterday
<i>an-dràsta</i>	now (non-contrastive)	<i>a' bhòn-dè</i>	day before yesterday
<i>a-nis(d)</i>	now (contrastive)	<i>an-uiridh</i>	last year
<i>an-diugh</i>	today	<i>an-raoir</i>	last night
<i>a-nochd</i>	tonight	<i>a-bhòn-raoir</i>	two nights ago
<i>am bliadhna</i>	this year	<i>mar-thà;</i>	already
		<i>mu-thràth</i>	

FUTURE		INDEFINITE	
<i>an eara(i)r</i>	day after tomorrow	<i>an còmhnaidh</i>	always
<i>a-màireach</i>	tomorrow	<i>greis</i>	for a while
<i>gu sìorraidh</i>	for eternity	<i>uaireannan</i>	sometimes
<i>an ath-oidhch'</i>	tomorrow night	<i>uaireigin</i>	sometime
<i>an ath-bhliadhna</i>	next year	<i>pulla (N. Uist)</i>	for a while
<i>tuilleadh</i>	for ever	<i>fhathast</i>	still, yet

2.2.5.3 Direction and Location

As mentioned in A§2.2.2.2, a contrast is often made in Gaelic between the marking of location and movement:

English	ScG Movement	ScG Location
up	<i>suas</i>	<i>shuas</i>
down	<i>sìos</i>	<i>shìos</i>
over here	<i>a-nall</i>	<i>a-bhos</i>
over there	<i>a-null</i>	<i>thall</i>

There is also an adverb for expressing movement away from up or down, either *a-nuas* or *a-nìos* (neutralising an older, now rarely maintained distinction between the two) depending on the dialect⁶⁴: e.g. *thig a-nuas* 'come down' (e.g. from upstairs).

The directions of the compass are as follows: *tuath* 'north'; *deas* 'south'; *ear* 'east'; *iar* 'west'. There is some variation in these forms depending on whether location, movement, or adjectival force is communicated. See Calder (1923) for more information on this subject and for a list of the large number of prepositions used to communicate direction and location.

2.2.5.4 Degree

There are a number of adverbs for the purposes of amplification, emphasising, downtoning, and hedging:

- a) amplifiers: e.g. *gu lèir* 'completely'; *gu mòr* 'greatly'; *direach* 'exactly'; *glè* 'very'; *ro* 'extremely'
- b) downtoners: e.g. *gu ìre* 'to an extent'; *gu ìre bheag* 'to some extent'; *an ìre mhath* 'to a large extent'
- c) emphatics: e.g. *gu dearbh* 'indeed'; *gu deimhinne* 'with certainty'; *gun teagamh* 'without a doubt'
- d) hedges: e.g. *cha mhòr* 'almost'; *'s dòcha* 'perhaps'; *math dh'fhaodte* 'perhaps'

Gaelic has no epistemic adverbs as such, which code the reported source of a clause's information (e.g. hearsay *versus* first hand observation). However, it does have a few evidential adverbs, such as those emphatics and hedges found above.

2.2.6 Prepositions

Gaelic has three classes of prepositions:

- *simple prepositions*, most of which can incorporate pronominal elements (then called 'prepositional pronouns')
- *compound prepositions*, consisting of an adjective, adverb, or noun + a simple preposition
- *complex prepositions*, which consist of a simple preposition + noun.

2.2.6.1 Simple Prepositions

Most simple prepositions are mono-syllabic and take the dative case (D) although a few take nominative/accusative (N) or genitive (G) case. Some cause lenition (^b) of the following noun:

⁶⁴ In Irish, this distinction is not neutralised: *anuas* means 'from above' while *anìos* means 'from below'.

Preposition	Gloss	Case it Traditionally Governs
<i>a^L</i>	'to' (reduced form of do)	D
<i>à, às</i>	'out of'	D
<i>aig</i>	'at'	D
<i>air</i>	'on'	D
<i>an, ann an</i>	'in'	D
<i>(bh)o^L</i>	'from'	D
<i>de^L</i>	'of' (partitive)	D
<i>do^L</i>	'to'/'for'	D
<i>eadar</i>	'between'	N
<i>far</i>	'off'	G
<i>fo^L</i>	'under'	D
<i>gu</i>	'to (the point of)'	D (G if definite)
<i>gun^L</i>	'without'	N
<i>le</i>	'with'	D
<i>mar^L</i>	'as'/'like'	D ⁶⁵
<i>mu^L</i>	'about'	D
<i>rè</i>	'during'	G
<i>ri</i>	'to'/'against'/'with'	D
<i>ro(imh)^L</i>	'before'	D
<i>tarsainn</i>	'across'	G
<i>thar</i>	'over'	G
<i>tro(imh)^L</i>	'through'	D

Of these, some can incorporate the definite article in one of two ways:

1) With a nasal /n/

bhon
gun/ chun < gu
d(h)en
don/ dhan < do
fon
mun
ron
tron

2) With a voiceless alveolar /s/ or postalveolar fricative /ʃ/

anns
leis
ris

The definite article may be optionally (group 1 above) or obligatorily reduplicated (group 2): *fon taigh/ fon an taigh* 'beneath the house'; *a's an taigh* Lit. 'in the the house'.

⁶⁵ Assigns nominative case when definite.

Prepositional Pronouns

Most of the simple prepositions can be inflected for person, gender, and number.

Table 135: Prepositional pronouns

<i>Preposition</i>	<i>Singular</i>				<i>Plural</i>		
	<i>1s</i>	<i>2s</i>	<i>3s-M</i>	<i>3s-F</i>	<i>1p</i>	<i>2p</i>	<i>3p</i>
<i>aig at</i>	agam	agad	aige	aice	againn	agaibh	aca
<i>air on</i>	orm	ort	air	oirre	oirnn	oirbh	orra
<i>ann in</i>	annam	annad	ann	innte	annainn	annaibh	annta
<i>às out of</i>	asam	asad	às	aisde	asainn	asaibh	asda
<i>(bh)o from</i>	(bh)uam	(bh)uat	(bh)uaithe	(bh)uaipe	(bh)uainn	(bh)uaibh	(bh)uapa
<i>de off</i>	dhiom	dhiot	dheth	dhiith	dhinn	dhìbh	dhiubh
<i>do for, to</i>	dhomh	dhut	dha	dhi	dhuinn	dhuibh	dhiubh
<i>eadar between</i>	--	--	--	--	eadarainn	eadaraibh	eadorra
<i>fo under</i>	fodham	fodhad	fodha	foidhpe	fodhainn	fodhaibh	fodhpa
<i>gu to</i>	(th)ugam	(th)ugad	(th)uige	(th)uice	(th)ugainn	(th)ugaibh	(th)uca
<i>le with</i>	leam	leat	leis	leatha	leinn	leibh	leotha
<i>mu about</i>	umam	umad	uime	uimpe	umainn	umaibh	umpa
<i>ri to, with, against</i>	rium	riut	ris	rithe	rinn	ribh	riutha
<i>ro(imh) before</i>	romham	romhad	roimhe	roimhpe	romhainn	romhaibh	romhpa
<i>tro(imh) through</i>	tromham	tromhad	troimhe	troimhpe	tromhainn	tromhaibh	tromhpa
<i>thar over</i>	tharam	tharad	thairis	thairte	tharainn	tharaibh	tharta
<i>Emphatic suffixes</i>	-sa	-sa	-san	-se	-c	-se	-san

Emphatic suffixes can be added to each of these forms, e.g. *agamsa*, *agadsa*, *aigesan*, *aicese*, *againne*, *agaibhse*, *acasan*.

Possessive Prepositional Pronouns

A small number of prepositions have fused with the possessive pronouns: *aig*, *ann*, *do*, and *ri*.⁶⁶ Instead of using **ann mo làimh* 'in my hand' for instance, '*nam làimh* or '*na mo làimh* is employed. These are presented in Table 136:

⁶⁶ In the Uist dialect at least, *air do* 'on your' may become *orra*.

Table 136: Possessive prepositional PNs

<i>Preposition</i>	<i>Singular Pronouns</i>				<i>Plural Pronouns</i>		
	<i>1s: mo^L</i>	<i>2s: do^L</i>	<i>3s-M: a^L</i>	<i>3s-F: a</i>	<i>1p: ar</i>	<i>2p: ur</i>	<i>3p: an/am</i>
<i>aig at</i>	gam ^L	gad ^L	ga ^L	ga	gar	gur	gan/gam
<i>ann in</i>	ga mo ^L	ga do ^L					
	'nam ^L	'nad ^L	'na ^L	'na	'nar	'nur	'nan/'nam
<i>do to/for</i>	'na mo ^L	'na do ^L					
	dham ^L	dhad ^L	dha ^L	dha	dhar	dhur	dhan/dham
<i>ri to</i>	dha mo ^L	dha do ^L					
	rim ^L	rid ^L	ri ^L	ri	ri ar	ri ur	rin/rim
	ri mo ^L	ri do ^L					

As the possessive pronoun is integral to these expressions, their use with vowels parallels that covered in A§2.2.1.5, for example: *gar n-èiridh* 'at our arising'.

2.2.6.2 Compound Prepositions: ADV, N, or ADJ + PREP

These always take dative case and may incorporate pronominal elements as above:

<i>barrachd air</i>	'in addition to', 'more than'
<i>ceangailte ri</i>	'tied to'
<i>an coimeas ri</i>	'compared to'
<i>coltach ri</i>	'similar to'
<i>còmhla ri/ cuide ri</i>	'along with'
<i>faig air</i>	'close to'
<i>goirid air</i>	'close to'
<i>làmh ri</i>	'next to', 'compared to'
<i>maille ri</i>	'along with'
<i>seachad air</i>	'past'
<i>suas ri</i>	'up to'
<i>an taca ri</i>	'next to', 'compared to'
<i>taobh ri</i>	'side to'
<i>thairis/ fairis air</i>	'over'
<i>timcheall air</i>	'about'

2.2.6.3 Complex Prepositions: PREP + N

As a noun is the terminal element in these expressions, they govern the genitive (except *gu ruige*):

<i>a dh'ionnsaigh</i>	'towards'	<i>am fianais</i>	'in sight of'
<i>a rèir</i>	'according to'	<i>am measg</i>	'in the midst of'
<i>a thaobh</i>	'about', 'regarding'	<i>an àite</i>	'in place of'
<i>air bheulaibh</i> ⁶⁷	'in front of'	<i>an dèidh</i>	'after'
<i>air chùl(aibh)</i>	'behind'	<i>as aonais</i>	'deprived of'
<i>air feadh</i>	'through', 'amongst'	<i>gu ruige (+ N)</i>	'to'
<i>air muin</i>	'on the back of'	<i>mu choinneimh</i>	'opposite'
<i>air sgàth</i>	'for the sake of'	<i>mu dheidhinn</i>	'regarding'
<i>air son</i>	'for the sake of'	<i>mu thimcheall</i>	'about'
<i>air tòir</i>	'in pursuit of'	<i>os cionn</i>	'above'
<i>an cois</i>	'near', 'with'	<i>ri linn</i>	'because of', 'during'

⁶⁷ -(a)ibh is the old dative plural ending

Complex prepositions are an obvious case of fusion in the language (see Hopper and Traugott 1993: 32-49) where two once separate forms – a preposition and a noun in the dative – have been reanalysed into a new compound:

cf. [[*air*] *chùlaibh an taighe*] > [[*air chùlaibh*] *an taighe*]
 [[back] of the house] > [back of [the house]]

However, there can still be 'opacity', ambiguity which in some contexts allows the compound to be analysed as before, e.g. the way that complex prepositions infix possessive pronouns:

air mo chùlaibh

on my back

NB: 'behind me', *not* 'on my back' which is *air mo dhruim*

2.2.7 Nominal Categorisers: Nominalisation and Diminutives

With the hybrid status of verbal nouns, ScG has no need for morphological operations to deal with action nominalisations such as 'swimming is good for you':

(175) *tha* *snàmh* *math* *dhut*
 be-PRES swimming good for you
 'swimming is good for you'

Compare the following:

(176) *tha* *thu* *a'* *snàmh* *gu* *math*
 be-PRES 2S PROG swimming-VN ADV good
 'you are swimming well'

However, there are several productive suffixes for denoting participant nominalisation. These include *-ar* (*-air(e)/-eir*), *-aiche*, *-adair*: e.g., *fastaidhear* 'employer' (< *fastadh* 'employ'); *snàmhache* 'swimmer' (< *snàmh* 'swim'); *riaghladair* 'ruler' (< *riaghladh* 'ruling').

While there are no productive augmentative suffixes in ScG, there are a few semi-productive diminutive ones. These occur especially in personal names and names of physical objects. Many nouns occurring with these suffixes are now fixed forms; some of the non-diminutive forms are non-extant:

Masculine names and nouns: *-an/lean*; *-agan*

Exaples: *fearan* 'wee man' (*fear* 'man'); *cnocan* 'hillock' (*cnoc* 'hill'); *Dòmhnallan* 'Donny' (*Dòmhnall* 'Donald'); *Uilleagan* 'Willie' (*Uilleam* 'William'); *cuilean* 'puppy'

Feminine names and nouns: *-ag*

Examples: *duanag* 'ditty' (*duan* 'poem'); *Mòrag* 'Sarah' (< *Mòr*); *Curstag* 'Kirsty' (< *Curstaidh*); *pis eag* 'kitten'

2.3 Verbal Morphology

While Gaelic is basically a VSO language, the initial verb is not always the predicating element. Depending on tense-aspect-modality parameters, it is sometimes the verbal noun which establishes lexical meaning while the initial verb, a sort of pro-verb or auxiliary, serves mainly to code tense, mood, and the absence or presence of negation. This can be seen in the following example, demonstrating the difference between the 'simple past' and 'progressive past':

(177) bha mi a' dol dhan bhùth(aidh)
 be-PAST 1S PROG go-VN to-ART shop-DAT
 'I was going to the shop'

(178) chaidh mi dhan bhùth(aidh)
 go-PAST 1S to-ART shop-DAT
 'I went to the shop'

The grammatical categories central to the ScG verbal system include tense, aspect, modality, voice, person and number. There are contrasts to be seen, as above, between inflected and periphrastic forms and, as a whole, periphrasis is more productive.

2.3.1 Verbs

The ScG verbal system is somewhat complex and only the main points can be covered here. For fuller accounts, the grammars in the bibliography should be consulted.

a) Distribution of the Verb

The finite verb always comes first in an utterance, unless it is preceded by particles (or 'sentence class markers' (Macaulay 1992)) marking negation, illocutionary force and interclausal relations. In transitive, non-periphrastic clauses, the verb is followed immediately by its two arguments (SUBJECT, OBJECT) as in the following:

V	S	O
<i>chunnaic</i>	<i>Ealasaid</i>	<i>Dùghall</i>
see-PAST	Elizabeth-N	Dugald-N
'Elizabeth saw Dugald'		

In periphrastic cases, the order of the first two elements remains the same, but the position of the second argument in relation to the verbal noun can vary depending on which aspectual/modal prepositions are employed:

V	S		VN	O
<i>tha</i>	<i>Ealasaid</i>	<i>a'</i>	<i>faicinn</i>	<i>Dhùghaill</i>
be-PRES	Elizabeth-N	PROG	see-VN	Dugald-G
'Elizabeth is seeing/sees Dugald'				

V		S		O	VN
<i>chan</i>	<i>cil</i>	<i>Ealasaid</i>	<i>air</i>	<i>Dùghall</i>	<i>fhaicinn</i>
NEG	be-PRES-DEP	Elizabeth-N	PERF	Dugald-N	seeing-VN
'Elizabeth is not after Dugald seeing': i.e., 'Elizabeth has not seen Dugald'					

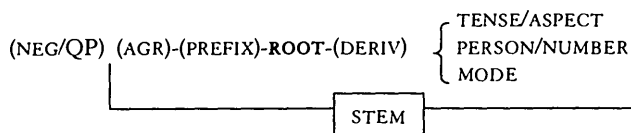
If one takes the semantic element to be the 'verb' (i.e. the verbal noun) rather than the finite one (see arguments in Cram 198), it is possible to interpret these two examples as instances of SVO and SOV order respectively. However, as there is always an obligatory verbal element in the first position in Scottish Gaelic, it is misleading to construe the basic syntax as anything other than VSO.

b) Structure of the Verb

The structure of the verb form depends on whether the verb in question is regular or irregular. For regular or 'weak' verbs, the same root – essentially the form of the 2nd person imperative – is present in all conditions. However, the ten irregular or 'strong' verbs feature suppletion and their stems are anomalously differentiated according to tense and the presence or absence of certain pre-verbs and sentence class markers. The

independent form of a verb is used in the absence of these pre-verbs and class markers. The dependent form is used when they are present. As will be seen in the following sections, there are various suffixes used for marking the tense-aspect-modality system and also a small number of person/number suffixes which interact with it.

The structure of regular, finite verbs can be represented as follows:



Regular verb: *òl* 'drink'

INDEPENDENT

*dh'òl e*⁶⁸ 'he drank'

òlaidh e 'he will drink'

dh'òladh e 'he would drink'

DEPENDENT

an do dh'òl e 'did he drink'

chan òl e 'he will not drink'

nan òladh e 'if he would drink'

Irregular verb: *rach* 'go'

INDEPENDENT

chaidh e 'he went'

thèid e 'he will go'

rachadh e 'he would go'

DEPENDENT

an deach e 'did he go'

cha tèid e 'he will not go'

nan rachadh e 'if he would go'

Amongst the pre-verbal particles/sentence class markers – affecting whether it assumes independent or dependent form – are the following:

Table 137: Pre-verbal particles

<i>Independent</i>	<i>Dependent</i>
<i>a</i> relativiser	<i>an/am</i> inter. clause marker
<i>ma</i> introduces conditional clause	<i>cha(n)</i> clausal negation
<i>na</i> introduces headless relative clauses	<i>gun/gum</i> complementiser
	<i>nachneg.</i> relativiser, neg. inter., neg. complementiser
	<i>nan/nam</i> conditional clause (with DEF PAST and INDEF2)

Some of these cause lenition in the following verb: *cha* always lenites lenitable consonants while *nach*, *an/am*, and *gun/gum* (depending on dialect) only lenite 'f'.

c) The Verbal Noun: Structure and Distribution

The verbal noun is generally comprised either by the root form itself or the root and a small number of suffixes such as *-ail*, *-aich*, *-amh*, *-inn*, and the most productive of them: *-adh*. Its structure can be represented as: (PREFIX)-ROOT-(SUFFIX). As mentioned before, the VN can function both as a noun and as a lexical verb. Its distribution, rather than morphology, is the only clue to its function:

⁶⁸ *Dh* is an affix used with vowel-initial verbs in lieu of lenition for certain tenses. See A§2.3.1.2.

- (179) tha e ag òl fiona
 be-PRES 3S-M PROG drinking-VN wine-G
 'he is drinking wine'
- (180) tha òl fiona ga mharbhadh
 be-PRES drinking-N wine-G PROG-3s-M-POSS killing-VN
 'wine drinking is killing him'
- (181) tha e 'dol a dh'òl fiona a-nochd
 be-PRES 3S-M PROG-going-VN drinking-INF wine-G tonight
 'he is going to drink wine tonight'

As is seen, there is no morphological difference between the three examples. Perhaps the best way to conceptualise the verbal noun, in essence, is as a noun which is: not time-stable in the way other nouns are; which can serve as a complement to another VN as in (181) above; and also be dominated by a small set of prepositions which convey aspectual meaning (179). Although the verbal nouns have been glossed and labelled here according to their functions, this is not meant to imply that the same form, as in *òl* above, requires specification three different times in the lexicon.

d) Verb classes

Macaulay (1992) states that Gaelic verbs are inherently either dynamic or stative but does not provide evidence for this breakdown. It is clear, however, that this distinction is morphosyntactically maintained with certain verbs that can assume *either* a static or dynamic interpretation. For instance, consider the way that Gaelic expresses 'stretched [out]' (state) *versus* 'stretching' (movement):

- (182) tha mi nam shineadh
 be-PRES I in my stretching
 'I am stretched out' (state) [as in lying down]
- (183) tha mi a' sineadh
 be-PRES I at stretching
 'I am stretching' (movement)

In English, the grammatical (usually human/animate) subject plays a large syntactic role but in ScG, as in the older stages of some other Indo-European languages, greater emphasis is placed on whether a participant is an AGENT or NON-AGENT.⁶⁹ Emotional expression is one of the areas in which this can be seen. In ScG, the EXPERIENCER of an emotion usually is said to have it 'on' or 'at' him/her. For instance, to say 'I am angry', the normal way is *tha fearg orm*, Lit. 'anger is on me'. To express 'she angered me', one would say *chuir i fearg orm*, Lit. 'she put anger on me'. Jealousy is expressed similarly: *tha farmad agam ort* 'there is jealousy at me on you' ~ 'I am jealous of you'.

Bodily functioning is managed similarly. Usually, the function itself is a noun and the verb activating it is of a general sort such as 'doing', 'letting', 'putting' or simply 'being'. For example:

- | | | | | | |
|-----------------|------------|-----------|-----------|----------------|-----------------|
| 'I am bleeding' | <i>tha</i> | <i>mi</i> | <i>a'</i> | <i>sileadh</i> | <i>fala(dh)</i> |
| | be-PRES | 1S | at | dripping-VN | blood-G |

⁶⁹ Thanks to Professor Jim Miller for this point.

'he sneezed'	<i>rinn</i>	<i>e</i>	<i>sreothart</i>
	do-PAST	3S-M	a sneeze
'she hiccuped'	<i>rinn</i>	<i>i</i>	<i>aileag</i>
	do-PAST	3S-F	hiccup
'she had the hiccups'	<i>bha</i>	<i>an aileag oirre</i>	
	be-PAST	ART hiccup	on her
'you vomited'	<i>chuir</i>	<i>thu a-mach</i>	
	put-PAST	2S	out
'they laughed'	<i>rinn</i>	<i>iad gàire</i>	
	do-PAST	3P	laugh

Weather, sensation/perception, cognition and manipulation are other verb classes that we could highlight in this regard. Available space, however, precludes further examination of this interesting area.

2.3.1.1 Personal Affixes

In Finite Verbs

Only in the 2nd indefinite ('conditional/habitual') and imperative are person and number normally marked on finite verbs in ScG. This occurs as a suffix attached directly to the verbal root. In the 2nd indefinite, only the 1S is now regularly marked in this way although some dialects preserve the 1P form. Using the regular verb *cuir* 'put' we can illustrate the pattern:

2nd indefinite: *chuirinn* 'I would put'; *chuireamaid* 'we would put' (conservative use) ~ *chuireadh sinn* 'we would put' (progressive use); cf. *chuireadh el/sibh/iad*

The full paradigm for the imperative is as follows although most younger speakers would only use the 2S form, and possibly the 2P one:

<i>cuirim</i>	let me put
<i>cuir (thusa)</i>	you put
<i>cuireadh (esan/ise)</i>	let him put
<i>cuireamaid</i>	let us put
<i>cuiribh</i>	let you (pl.) put
<i>cuireadh (iadsan)</i>	let them put

Very rarely, 'present' tense forms occur for the 1st person singular, particularly with psychological verbs. See A§2.1.1 for examples.

In Non-finite Verbs

When a verbal noun takes a pronominal object, it is incorporated in the form of a possessive prepositional pronoun, functionally a pronominalised aspect marker:

- (184) *tha mi ga (aig 'at' + a 'his/its') thuigsinn*
 be-PRES 1S PROG-3S-POSS understanding-VN
 'I understand him/it' [Lit. 'I am at his/ its understanding']

Using the same verb and progressive aspect, here is a chart of all of these object pronouns:

<i>gam thuigsinn</i>	understanding	me
<i>gad thuigsinn</i>	...	you (sing.)
<i>ga thuigsinn</i>	...	him
<i>ga tuigsinn</i>	...	her
<i>gar thuigsinn</i>	...	us
<i>gur thuigsinn</i>	...	you (pl./polite)
<i>gan tuigsinn</i>	...	them

Objects are incorporated into infinitives in a similar way but they are spelled and pronounced slightly differently. In essence, the progressive of the verbal noun is constructed by *ag* 'at' (< *aig* prep.) while the infinitive is marked by *a* ^(L) which is a short form of the preposition *do* ^(L) 'to/for':

(185)	tha	mi	'dol	dha	thuigsinn
	be-PRES	1S	PROG-doing-VN	to its	understanding-INF
	'I am going to understand it'				

Impersonal Affixes

Scottish Gaelic has a set of impersonal suffixes which can be attached to the verbal root. These are differentiated by tense and whether a verb is strong or weak (irregular or regular). The suffix *-ar*⁷⁰ is used with the present tense (which only occurs with *bi* 'to be') and with all 1st indefinite forms (the 'future' tense). It may also be used with the past tense of the verb 'to be': *bhathar*. The ending *-adh* is used with the past forms of weak verbs while *-as* is used with strong verbs. Finally, *-te/-ta/-ist(e)* is used with all 2nd indefinite forms (the conditional/habitual). See Appendix 1 for a chart of all of the strong verb forms. The following illustrates these endings with the strong verb *bi* (other strong verbs are formed similarly), and the weak verb *cuir*:

INDEPENDENT	GLOSS	DEPENDENT	GLOSS
<i>thathar ga dhèanamh</i>	it is being done	<i>a bheilear ga dhèanamh</i>	is it being done
<i>bithear ...</i>	it will be done	<i>am bithear ...</i>	will it be done
<i>bhathas ...</i>	it was being done	<i>a robhas ...</i>	was it being done
<i>bhite ...</i>	it would be done	<i>am bite ...</i>	would it be done

INDEPENDENT	GLOSS	DEPENDENT	GLOSS
<i>cuirear e</i>	it will be put	<i>an cuirear e</i>	will it be put
<i>chuireadh e</i>	it was put	<i>an cuireadh e</i>	was it put
<i>chuirte e</i>	it would be put	<i>an cuirte e</i>	would it be put

2.3.1.2 Tense, Aspect and Mode

As in many other languages, tense, aspect and mode are not always clearly differentiated in Gaelic. Temporal interpretation can be quite flexible, especially when an utterance is strongly irrealis in mode.

Tense

As Macaulay⁷¹ (1996) points out, many grammarians working on Scottish Gaelic have been too ready to impose the tense systems of Classical languages, Classical Gaelic and even

⁷⁰ The ending *-as* is sometimes used as well in present and past tenses, e.g. *chunnacas e a' tighinn* 'he was seen coming' or *thathas ga dhèanamh* 'it is being done'. There is a distinction in some dialects between events implicating human presence or action (*-ar*) and those without this quality (*-as*): an 'impersonal' impersonal.

⁷¹ One of the few native ScG-speaking linguists working on the language.

Modern English upon ScG without addressing its own inherent contrasts. He proposes a paradigm based upon oppositions between tensed and non-tensed on the one hand, and definite and indefinite 'mode' on the other. He maintains (1992) that there are only two tenses: **present** ('non-past' in Macaulay 1996) and **past**, which are both definite in modality. The untensed, indefinite mode is broken down into the '**first indefinite**' – a kind of flimsy future – and the '**second indefinite**' which is similarly flexible and deals generally with conditional and habitual frames of reference. This leads to a more sympathetic account of the facts. In particular it addresses the tendency in the language to rely on context and extraclausal periphrastic devices (e.g. adverbials) as opposed to inflection for the placing of events in relation to the time of speaking.

Scottish Gaelic Tense

DEFINITE MODE (Tensed)		INDEFINITE MODE (Untensed)	
past	non-past	Indef 1	Indef 2

The verb *bi* 'to be', often used as an auxiliary in conjunction with a verbal noun (see (181)), is the only verb with a 'definite non-past' ('present tense') form. It is inflected in Macaulay's paradigm as follows (adapted from 1992;1996):

Root: *bi*

DEFINITE

PRES	INDEP	<i>tha</i>	'is'
	DEP	<i>bheil</i>	
PAST	INDEP	<i>bha</i>	'was'
	DEP	<i>robh</i>	

INDEFINITE

1 ST	INDEP	<i>bidh/bithidh</i>	'is/will be'
	REL	<i>bhios</i>	
	DEP	<i>b(h)i</i>	
2 ND	INDEP	<i>bhitheadh/bhiodh</i>	'would be/ used to be'
	DEP	<i>bitheadh/biodh</i>	

The strong verbs (such as *faic* 'to see') are inflected as above except, as mentioned, they do not have a definite-non past (see Appendix 1). They also do not have a distinct relative form (often called the 'future-relative') in INDEF1. Weak verbs differ from strong verbs in not having distinct dependent forms. They *do* however have a distinct relative form in INDEF1. Inflectional morphemes and affixes mark the different reflexes of the tense system, summarised in Table 138:

Table 138: Weak verb inflection

	PAST _{con}	PAST _{vow}	INDEF1 _{con}	INDEF1 _{vow}	INDEF2 _{con}	INDEF2 _{vow}
INDEP	^L stem	<i>dh'</i> + stem	stem + <i>-(a)idh</i>	stem + <i>-(a)idh</i>	^L stem + suffix (see A §2.3.1.1)	<i>dh'</i> + stem + suffix
DEP	<i>do</i> + ^L stem	<i>do dh'</i> + stem	(^L)stem	stem	(^L)stem + suffix	stem + suffix

Here, 'con' refers to consonant-initial verbs and 'vow' refers to vowel-initial verbs. Verbs beginning with 'f' receive the *dh'* affix as vowel-initial verbs but lenite as consonants in those situations mentioned in A§2.3.1b above. When a verb begins with a 'non-lenitable' consonant in a lenitable situation, no orthographic change is signified although there may be changes in pronunciation (i.e. with the sonorants: l; n; r). Orthographically, suffixes are assimilated to the stem according to the broad-broad and slender-slender rules and whether or not a dependent form lenites is determined by the particle preceding it.

An illustration of the paradigm for a consonant-initial verb and a vowel-initial verb follows:

Root: *coisich* 'walk'

DEFINITE

PAST	{ INDEP <i>choisich</i>		'walked'
	DEP	<i>do choisich</i>	

INDEFINITE

1 ST	{ INDEP <i>choisichidh</i>		'will walk'
	REL	<i>choisicheas</i>	
	DEP	<i>coisich</i>	
2 ND	{ INDEP <i>choisicheadh</i> , etc.		'would walk/used to walk'
	DEP	<i>coisicheadh</i> , etc.	

Root: *òl* 'drink'

DEFINITE

PAST	{ INDEP <i>dh'òl</i>		'drank'
	DEP	<i>do dh'òl</i>	

INDEFINITE

1 ST	{ INDEP <i>òlaidh</i>		'will drink'
	REL	<i>dh'òlas</i>	
	DEP	<i>òl</i>	
2 ND	{ INDEP <i>dh'òladh</i> , etc.		'would drink/used to drink'
	DEP	<i>òladh</i> , etc.	

Macaulay's labels for the indefinite mode (1st and 2nd) are deliberately opaque; he is attempting to side-step the traditional names, which he considers inappropriate. In traditional grammars, the 1st indefinite is called the 'present-future' while the 2nd indefinite is usually called the 'past-habitual'. A closer look at the indefinite mode reveals that is actually more ambiguous than these labels would suggest (adapted from Macaulay 1992: 219):

1ST INDEFINITE

- (186) *bidh e sa chàr a-nèisd*⁷²
 be-INDEF1 3S-M in-ART car-D now-CONTR
 'he'll be/is in the car *now*' (i.e. he wasn't before)

- (187) *bidh e ann a-màireach*
 be-INDEF1 3S-M 'in it' tomorrow
 'he'll be there/here tomorrow'

⁷² *A-nèisd* is a dialectal form of *a-nis*, as used in North Uist.

- (188) bidh e ann a h-uile h-oidhche
 be-INDEF1 3S-M 'in it' every night
 'he is/will be there/here every night'
- (189) bidh e bochd is dòcha
 be-INDEF1 3S-M ill perhaps
 'he will be/is ill perhaps'

The first example (186) is the kind of clause that is often used to support calling this the 'present-future' as it crosses into both tenses. However, it clearly lacks the conviction of *tha e sa chàr a-neis* 'he is in the car now' which is a definite, declarative statement referring to the 'true present'. The second (187) has future reference, yet (188) is iterative and (189) is speculative. It is clear that *bidh* differs from *tha* in having a less defined sense of 'realis', a common characteristic of future tense, which by definition has yet to occur. This characteristic is part of what Macaulay is trying to convey by his label 'first indefinite'.

The first indefinite is often used to establish ability:

- (190) nì mi sin gun trioblaid sam bith
 do-INDEF1 1S that without trouble at all
 'I can do/will do that with no trouble at all'

Also, the first indefinite is commonly used with verbs of perception to refer to the 'true present':

- (191) an cluinn thu an ceòl sin
 INT hear-INDEF1-DEP 2S ART music that
 'do you hear that music?'

Finally, the verbal noun *dol* 'going' used in conjunction with an infinitive can be used to convey future time:

- (192) tha mi 'dol a choiseachd dhachaigh a-màireach
 be-PRES 1s PROG-going-VN to walking-INF home tomorrow
 'I am going to walk home tomorrow'

The second indefinite is ambiguous in a similar fashion. For example, one can construe the following as meaning 'he would always be there' both in the 'habitual-past' (every Thursday) and 'conditional-future' (if he received steady remuneration):

- (193) bhiodh e ann an-còmhnaidh
 be-INDEF2 he in it always
 'he would always be there/here'

In essence, it is possible to demonstrate that these verb forms do not grammaticalise tense in the way *tha* and *bha* do; they open to a wide range of interpretations depending on context and peripheral modifiers. The interested reader is referred to the references above for more information. Macaulay himself says that his suggestions are not meant to be conclusive and much more could be said about this area.

Aspect

There are three main categories of aspect in ScG: progressive, signifying a continuing, dynamic process; perfect⁷³, expressing a state in the 'present'⁷⁴ resulting from an earlier situation (usually an activity) as denoted by the verb; and prospective, indicating that something is going to occur. These are constructed periphrastically using aspectual particles⁷⁵ in coordination with an auxiliary verb (all the forms of *bi* are possible) and verbal noun. They can also combine to create compound aspect.

Progressive aspect is marked by the particle *ag* or *a'* meaning 'at':

- (194) *tha* *iad* *a'* *togail* *na* *cloiche(adh)*
 be-PRES 3P PROG building-VN ART rock-G
 'they are lifting the rock' (Lit.: 'they are at the lifting of the rock')

Perfect aspect is expressed by the particle *air* 'after':

- (195) *tha* *iad* *air* *a'* *chlach* *a thogail*
 be-PRES 3P PERF ART rock lifting-VN
 'they have lifted the rock' (Lit.: 'they are after lifting the rock')

cf. *tha iad air a' chloich* (Dat.) 'they are on the rock'

Perfective aspect on the other hand would be expressed through the definite past form: e.g. *thog iad an taigh seo* 'they built (raised) this house'. While one could quite felicitously utter this clause in combination with ... *agus thog iad fear eile as a dhèidh sin* '... and they built another one after that', it seems pragmatically strange if combined with one having perfect aspect: ?*tha iad air an taigh seo a thogail agus thog iad fear eile as a dhèidh sin* 'they have built this house and they built another one after that'.

Prospective aspect is expressed in a similar manner to the perfect, using *gu(s)*:

- (196) *tha iad gu a' chlach a thogail* 'they are about to lift the rock'

These are commonly combined to form compound aspect, often using the infinitive form *a bhith* 'being':

- (197) *tha iad air a bhith a' togail na cloiche(adh)* (perfect + progressive)
 'they have been lifting the rock'
 (198) *tha iad air a bhith gu a' chlach a thogail* (perfect + prospective)
 'they have been about to lift the rock'
 (199) *tha iad gu bhith a' togail na cloich(eadh)* (prospective + progressive)
 'they are about to be lifting the rock'

Perfect aspect and past tense combine to form the pluperfect: :

- (200) *bha* *iad* *air* *a'* *chlach* *a* *thogail*
 be-DEF-PAST 3P PERF ART rock-N AGR lifting-VN
 'they had lifted the rock'

⁷³ Macaulay calls this 'perfective' but the perfective—being a "situation viewed in its entirety" (Payne 1997, 239)—is incompatible with progressive aspect, a combination which occurs in ScG.

⁷⁴ By 'present' is meant being co-temporal with the internal temporal structure of the situation expressed by a clause.

⁷⁵ These are clearly cognate with prepositions of similar form (i.e. *aig*, *air*, *ri*, *gu*) but they do not govern dative case and have acquired special aspectual connotations.

Other aspectual distinctions are realised periphrastically and often rely, as with the habitual, on peripheral modifiers to convey the particular interpretation desired:

- (201) *chì mi an duine sin a h-uile latha* 'I see that man every day'
(cf. *chì mi an duine sin a-neis* 'I (can) see that man now')

Mode

Because of the potential confusion created by the terms *mode* and *mood*, we will deconstruct the notion and refer to three different types: a) deontic modality; b) epistemic modality; and c) illocutionary force (Van Valin and LaPolla 1997). We will take *mode* to mean, in general, the way a speaker frames an utterance in terms of its degree and/or kind of reality. In ScG there are two modal verbs – *feum* (must, oblige) and *faod* (can, may) – both of which are inflected as weak verbs but they only have forms in the INDEF1 and INDEF2.

The distinction between realis and irrealis seems to be grammaticalised to some extent in ScG strong verbs with realis broadly realised by independent forms and irrealis by dependent forms. Along with interrogatives and imperatives, negation seems to be treated as irrealis as well. However, the dependent verbal form is also used in the complement clause (see A§3.3.2), which is immaterial to the issue, therefore the distinction is not fully delineated.

IRREALIS

Imperatives (root form): *dèan sin* 'do that'; *faigh sin* 'get that'; *faic* 'look'

Interrogatives and Negatives: *cha dèan thu sin* 'you won't do that'; *am faigh thu sin* 'will you get that'; *am faic mi e* 'will I see it?'

REALIS

Declarative: *nì mi sin* 'I'll do that'; *gheibh thu sin* 'you'll get that'; *chì iad e* 'they see/will see/can see it'

a) Deontic Modality

Deontic modality refers to cases of obligation, ability and permission: essentially a relation between the subject NP and an action. In Scottish Gaelic, there are several ways to express obligation:

feum + NPSUB + (NPOBJ) + INF

feumaidh mi dhol dhan bhùth(aidh) 'I must go to the shop'

bi + *aig* + (NPOBL) + *ri* 'to' + (NPOBJ) + INF/

bi + (NPSUB) + *aig* + (NP) + *ri* 'to' + INF

tha agam ri litir a sgrìobhadh 'I have to write a letter'

tha aig Seumas ri litir a sgrìobhadh 'James has to write a letter'

tha litir agam ri sgrìobhadh 'I have a letter to write'

tha litir aig Seumas ri sgrìobhadh 'James has a letter to write'

is/bu [the copula] + *c(h)òir* + *do* + (NP) + (NPOBJ) + INF

bu chòir dhut suidhe 'you ought to sit'

bu chòir do Bheathaig an obair a dhèanamh 'Betty should do the work'

is/b' + f(h)eudar + (as previous example)

is fheudar dhomh falbh 'I ought to leave'

Ability is conveyed generally using one of the following:

rach 'go' + *aig* + (NP) + *air* + (NPOBJ) + INF
chaidh againn air a bhuannachadh 'we managed to win it'

is/b' + *urrainn* + *do*⁷⁶ + (NP) + (NPOBJ) + INF
is urrainn dhuibh sin a dhèanamh 'you (pl.) can do that'

bi + *comasach* + *e* + *do* + (NP) + (NPOBJ) + INF
a bheil e comasach dhut bruidhinn 'is it possible for you to speak'

amaid 'aim, find' + NP + *air* + (NPOBJ) + INF
cha do dh'amaid e air a dhèanamh 'he did not manage to do it'

Permission is expressed using the verb *faod*:

<i>am faod mi snàmh?</i>	'may I swim?'
<i>faodaidh tu snàmh</i>	'you may swim'
<i>thuirt i gum faodainn falbh</i>	'she said that I was permitted to go'

b) Epistemic modality: necessity and possibility

As with similar verbs in English, there is some crossover in the way that one can interpret *feum* and *faod*. For instance, *feumaidh tu an càr a ghluasad far an rathaid* can mean both 'you are obliged to move the car off the road' (deontic reading: '...by order of the policeman') or 'it is necessary for you to...' (epistemic reading: '...in order for the other cars to get past'). Similarly *faodaidh iad buannachadh* could be construed as both 'they are able to win' (deontic reading) and 'it is possible they will win' (epistemic reading).

Hedges, such as *is dòcha*, (*math*) *dh'fhaodte*, and *is mathaid* are also used for expressing possibility with clausal complements (see also A§ 2.2.5.4 above):

(202) *dh'fhaodte gum faod sinn snàmh* 'it's possible that we might swim' (if the tide is out)

(203) *is dòcha gun tachair e fhathast* 'perhaps it will happen yet'

c) Illocutionary force is signalled in Gaelic primarily through morphology (e.g. lenition, suffixation, clitics, independent *versus* dependent verbal forms) and the presence of verbal particles. These particles (see A§2.3.1b), denoting interrogative utterances, negation, and interclausal dependency, present themselves before the finite verb.

Declaratives are denoted by the independent verb form and zero marking:

Strong verb: *bidh mi air n-ais* 'I'll be back'

Weak verb: *ithidh iad gu luath* 'they eat quickly'

Polar questions are expressed by using an interrogative particle in conjunction with the dependent form of the verb:

(204)	an	tèid	thu	leam	a Mhàiri
	QP	go-INDEF-DEP	2S	with-1S	Mary-V
	'will you go with me Mary?'				

⁷⁶ In some dialects, the subject of this construction can be a straight pronominal for some person/number categories: *is urrainn mi* 'I can'

Polar questions are also called 'yes/no' questions in the literature but affirmation and negation are not lexically grammaticalised in ScG. The closest word to a positive response is *seadh* meaning 'right/aye/go on', built from the copula and the old neuter pronoun *eadh* (cognate with Latin *id*). Affirmation is expressed by repeating the independent form of the focal verb from the question while negation is done by using *cha(n)* in conjunction with the dependent form of the verb:

<i>thèid</i>	'yes' [Lit. 'will go']
<i>cha tèid</i>	'no' [Lit. 'will not go']

Content questions are formed using one of the so called *wh*-question words (see A§2.2.2.4). When question words have oblique roles, usually adpositions are fronted (pied-piped) but they sometimes remain with the 'gap', especially in less formal registers and in spontaneous spoken language:

- (205) *cò ris a bha thu a' bruidhinn*
 who to REL be-PAST 2S PROG speaking-VN
 'to whom were you speaking?'

- (206) *cò a bha thu a' bruidhinn ris*
 'who were you speaking to?'

Tag questions come in three types: one with similar polarity⁷⁷ and two with opposite polarity. These take the form of elided interrogative clauses:

- (207) *tha thu a' falbh a bheil* [same polarity]
 be-PRES 2S PROG depart-VN QP be-PRES-DEP
 'you are leaving, are you?'
- (208) *tha thu a' falbh nach eil* [different polarity: leading question]
 'you are leaving, aren't you?'
- (209) *chan eil thu a' falbh a bheil* [different polarity: leading question]
 'you are not leaving, are you?'

Direct imperatives are formed using the verb stem and any personal suffixes required (see A§2.3.1.1). As mentioned before, only the 2nd person imperatives are very productive today:

- (210) *dùin do bheul* 'shut up' [Lit. 'shut your mouth']
 (211) *dèanamaid ùrnaigh* 'let us pray'

The positive and negative interrogative forms of the first indefinite can be used to communicate indirect imperatives:

- (212) *am bi thu sàmhach* 'will you be quiet'
 (213) *nach tig thu* 'won't you come'

⁷⁷ Macaulay (1992) also mentions the following, although I have never heard it in spoken language nor seen it in writing:

<i>chan eil</i>	<i>thu a'</i>	<i>falbh</i>	<i>nach</i>	<i>eil</i>	[same polarity]
NEG be-PRES-DEP	2S	PROG depart-VN	QP-NEG	be-PRES DEP	
'you are not leaving, you're not?'					

There are a few verbs with special imperative forms such as the verb 'to go': *thalla/fhalbh* 'go'; *thugainn/tiugainn* 'come'; *trobhad* 'come'; *theirig* 'go'.

2.3.1.3 Negation

Different parts of the clause can be negated in ScG. The pre-verbal particles *cha(n)* and *nach* serve as analytic clausal negatives:

- (214) chan eil iad gu math
 NEG be-PRES-DEP 3P well
 'they are not well'

- (215) nach eil iad gu math
 'are they not well?'

'Double negative' constructions are common in ScG, effectively cancelling each other out:

- (216) cha chreid mi nach eil iad gu math
 NEG believe-INDEF1 IS NEG-COMP be-PRES 3P well
 'I believe they are well' [Lit. I don't believe that they are not well]

Negative imperatives are expressed by placing the particle *na* before the verb:

- (217) na ithibh sin
 'don't (polite form) eat that'

Constituent negation is accomplished by the prepositions *gun* (+ NP_{NOM}) and *as aonais* (+ NP_{GEN}): *duine gun chiall* 'a man without sense'; *as ur n-aonais* 'without you (2P)'; *as aonais nan each* 'without the horses'. *Gun* can also be used to negate non-finite complements ('small clauses'):

- (218) bha i ag iarraidh orm gun na h-uighean ithe
 be-PAST she at want-VN on-1S neg ART eggs eat-INF
 'she wanted me not to eat the eggs'

- (219) cf. bha i ag iarraidh orm na h-uighean ithe
 be-PAST she at want-VN on-1S ART eggs eat-INF
 'she wanted me to eat the eggs'

Finally, there are various derivational negative prefixes such as *mì-*, *neo-*, *do-* and the generally unproductive *a(i)n-*, *ao-*, and *eu-*: *mì-nàdarrach* 'unnatural'; *neo-bhlasta* 'unpalatable'; *do-sheachanta* 'unavoidable'; *an-fhoiseil* 'restless'; *aotrom* 'light' [Lit. 'unheavy']; *eu-choltach* 'dissimilar'.

2.3.2 Verbal Categorisers

2.3.2.1 Valence Increasing Constructions: Causatives

As with most languages, there are lexical causatives in ScG, verbs where the notion of CAUSE is inherent, such as *bris(t)* 'break': *bhris Seumas an uinneag* 'James broke the window' [CAUSE(Seumas, BREAK(Seumas, window))]. There are also at least two pairs of semantically related verbs where one has a causal meaning and the other a non-causal meaning:

(224) Intransitive Caused Event
thug *Māiri* *air* [*Seumas* *ithe*] 'Mary made Seumas eat'
 1 2

(225) *thug Màiri [aṛ a' bhò a mharbhadh]* 'Mary made him kill the cow'
 1 2 3

(226) *Thàinig air [Seumas sin a dhèanamh]* 'Seumas was made to do that'

- they are semantically transitive (criterion 1)
- the participant that is best characterised as AGENT ('A') is either omitted or assumes an oblique role (criterion 2)
- the PASSIVE ('P') participant is in subject position (criterion 3)
- the verb has any and all of the characteristics of intransitive verbs in the given language (criterion 4)
- the discourse function of prototypical passive utterances is to highlight P in cases where it has more topicality than A (criterion 5)

There are four main constructions in ScG that function to highlight P over A. Each differs in terms of how it satisfies the other criteria above and there is no construction that is categorically ‘passive’ per se. We will take them in turn beginning with the one most like the prototype and ending with the one most unlike it.

Apparently, Gaelic has no lexical or morphological passives. However, like English, it has a periphrastic passive in the form of a 'past participle'. Their distribution is, as expected, similar to the predicative adjective (A§2.2.4) but they are not so productive. Morphologically, they are characterised by the ending *-te/ta*:

- (227) tha an sgeul sgrìobhte anns an leabhar
 be-PRES ART story-N write-PART in the ART book-D
 'the story is written in the book'

Perhaps a good way to construe the past-participle is as an adjectivised verb. Agents can be incorporated using the preposition *aig* 'at' or, patterned on the English, *le* 'with/by' which was traditionally reserved for instrumental rather than agentive uses (Macaulay 1993: 177):

- (228) bha an t-iomradh deasaichte aig/le Eàirdsidh
 be-PAST ART report-N prepare-PART at/by Archie-OBL
 'the report was prepared by Archie'
- (229) bha an t-iomradh deasaichte le compiutair
 'the report was prepared with a computer'

As seen, the lexical verb in these examples is semantically transitive (crit. 1 above), A has been demoted or omitted (crit. 2), P is in subject position (crit. 3), and the verb is morphosyntactically intransitive (crit. 4). This is an example of a personal passive in ScG.

b. Passives in *bi* + NP + ASP + (PNPOSS) + ⁽¹⁾VN

When an argument in subject position is co-referential with the pronominal object of an aspect-marker (*air*, *ag*, *ri*) the utterance takes on a passive interpretation:

- (230) tha an rìgh, ga, mharbhadh aig na searbhantan
 be-PRES ART king at-his killing-VN at ART servants
 'the king is being killed by the servants'
- (231) bidh an rìgh, air a, mharbhadh⁷⁸
 'the king will be killed'
 Lit. 'the king will be after his killing'

This expression fulfils criteria 1, 2, 3 and 5. However, as the verbal noun has an object, it cannot be construed as intransitive (crit. 4). Also, as ScG employs the same morphosyntax for the expression of aspect with pronominal objects in general, one cannot say that this construction is passive in form per se:

- (232) tha an rìgh, ga, mharbhadh
 'the king is killing him'

c. Passives in *rach* + NP/PN_{POSS} + ⁽¹⁾VN

These are characterised as 'non-agentive' expressions in Macaulay (1993: 177) where *rach* 'to go' is as an 'auxiliary' with the meaning of 'effect/come to pass'. Usually, this construction is used with semantically transitive lexical verbs:

- (233) thèid am ball a phutadh a-steach
 go-INDEFART ball AGR kicking-VN in-DIR
 'the ball will be kicked in'
- (234) thèid a phutadh a-steach
 'it will be kicked in'

⁷⁸ Payne (1997, 173) mentions the synchronic and diachronic relationship between passive voice and perfect aspect, seen clearly in this construction.

However, it can also be used felicitously with some semantically intransitive verbs (example and translation from Macaulay 1993: 177):

- (235) chaidh falbh
go-PAST departing-VN
'a departing was effected (?)'

When used with semantically transitive verbs, as it most frequently is, this construction fulfils all of our criteria for being an analytic impersonal passive. However, when it is used with intransitive verbs, it fails on criteria 1, 3, and 5.

d. Impersonal Endings

Touched on in A§2.3.1.1, ScG has a set of impersonal suffixes which can attach to the end of a finite verb. These are differentiated according to tense and whether a verb is weak or strong (see Appendix 1B for a full list). Clauses featuring them are often translated into English using its passive construction, but the French use of *on* is a more accurate parallel. In terms of morphosyntax, they are in complementary distribution to overt subjects occurring directly after the verb:

- (236) chreachadh am baile
destroy-IMPERS ART town-N
'the town was destroyed' ~ '(x) destroyed the town'
- (237) cf. *chreach iad a' bhaile* 'they destroyed the town'
- (238) *chreachadh iad a' bhaile* 'they would destroy the town'

The suffix is not a marker of the passive: an AGENT *is* marked, it is simply unspecified. As in the case of the imperative personal endings or those with INDEF2, it is a synthetic pronominal:

- chreachamaid am baile* 'we would destroy the town'
- creachamaid am baile* 'let us destroy the town'

Therefore, this construction is a case of active voice although, semantically, A is demoted and P is topicalised.

2.3.3 Verbal Modifiers: Participles and Gerunds

Strictly speaking, Scottish Gaelic does not have participles (Gillies 1993: 205), save the past-participle as mentioned. Neither does it have a clearly defined class of gerunds. The verbal noun is found where English and other languages would use participles and gerunds. Some examples will suffice:

- (239) chunnaic mi ag obair air stamh i
see-PAST 1S PROG work-VN on sea-tangle-D 3S-F
'I saw her working on sea-tangle'
- (240) tha a bhith 'g òl beagan fiona math dhut
be-PRES be-INF PROG drinking-VN little wine-G good for-2S
'drinking a little wine is good for you'

Because the verbal noun already has nominal status in most conditions, there is no need for additional morphology:

- (241) 's e fìor dhroch dhraibheadh a bha an-sin
 COP-PRES 3S true bad driving-N REL be-PAST there
 'that was truly bad driving'

3 Syntax

3.1 Predicate Nominals and Similar Constructions

There are two verbs 'to be' in Gaelic, *tha* (A§2.3.1.2) – once called the 'substantive' – and the defective copula *is*. The copula has only two inflections, the so-called 'present-future' and the 'past-conditional', each with an independent, a dependent, and a relative form:

<i>Tense</i>	<i>Independent</i>	<i>Dependent</i>	<i>Relative</i>
'Present-future'	is/s	gur/gun	As
'Past-conditional'	bu ^(1,) /b'	gum bu ^(1,) /b'	a bu ^(1,) /a b'

Essentially, both the definite non-past and indefinite 1 are collapsed in the 'present-future' form as are the definite past and indefinite 2 in the 'past-conditional'. The old opposition between the copula denoting long-lasting and/or inherent characteristics and *THA*, transient and/or superficial ones is largely defunct (Gillies 1993: 209–11) but its influence can still be detected in a limited number of ways as seen below.

3.1.1 Predicate Nominals: Proper Inclusion and Equation

a. COP + NP juxtaposition

Inalienable class membership was once formally denoted by the construction *COP* + *NP^{PREL}* + *NP^{SUB}* but it is now antiquated or poetic seeming:

- (242) is croitear Ùisdean
 COP-PRES crofter Hugh
 'Hugh is a crofter'

- (243) bu chroitear e 'he was/would be a crofter'

An equative construction obtains from the order *COP* (+ *e*) + *NP^{SUB}* + *DEF NP^{PREL}*. If the subject NP is pronominal, its emphatic form is used (A§2.2.2.1). If the subject NP is a proper noun, *is e* is always used. The predicate is always definite:

- (244) is (e) mise Raibeart 'I am Robert'
 (245) is (e) ise mo mhàthair 'she is my mother'
 (246) is e Iain am bodach sin 'Ian is that old man'

b. COP + 'e' + NP^{PREL} + *THA* + *ANN* (+ NP^{SUB})

The 'identificational cleft' is the most common way of expressing proper inclusion. The third person singular pronoun 'e' is frequently associated with the copular verb when used predicatively:

- (247) is⁷⁹ e croitear a th' ann an Ùisdean
 COP-PRES 3S crofter REL be-PRES in Hugh
 'Hugh is a crofter'

ANN is a prepositional pronoun, thus able to incorporate pronominal NPs:

⁷⁹ The copular verb is often omitted in speech in these kinds of expressions: *croitear a th' ann* 'he is a crofter'

- (248) is e croitear a th' annam 'I am a crofter'

The use of the 'past-conditional' form of the copula in cleft-type utterances has declined and, although it is still felt to be 'correct' to use it, tense marking is usually consigned to the relative verb, as in example (250) below:

- (249) b' e croitear a bh' ann 'he was a crofter' [Lit. 'it was a crofter that he was']
 (250) 's e croitear a bh' ann 'he was a crofter' [Lit. 'it is a crofter that he was']

c. *THA* + NP^{SUB} + *ANN*^{POSS} + NP^{PRED}

In some dialects, there still may be a contrast between this construction and the prior one in terms of perceived transience and inherence; *tha* may convey a certain measure of evanescence in relation to *is*, which is generally associated with permanence. However one of my older North Uist informants said they were synonymous to him.

Statement form

- (251) tha e 'na chroitear
 be-PRES IS 'in his' crofter
 'he is a crofter'

Question form

- (252) a bheil thu 'nad chroitear 'are you a crofter?'

Gillies (1993: 211) uses the following to highlight the potential difference in meaning that *b* and *c* can convey: *tha e 'na oileanach ach chan e oileanach a th' ann* 'he is (registered as) a student but he is not a student (by disposition)'.

3.1.2 Predicate Adjectives

These are distinct in form from predicate nominals in that they generally must use *THA* and not the copula (*THA* + NP^{SUB} + ADJ):

- (253) *tha Tòmas àrd* 'Thomas is tall'
 **is Tòmas àrd*

The form COP + ADJ + NP^{SUB} is found in older texts and poetic usage primarily, similarly to COP + NP^{PRED} + NP^{SUB} (see above). However it occurs in certain set phrases such as *is math sin* 'that is good', where 'that' is properly included in a set of good things whereas in *tha sin math*, 'that' is simply characterised in a transient sense as being 'good' (Gillies 1993). Another example:

- (254) bu bhrèagha a gnùis
 COP-PAST beautiful 3S-F visage
 'beautiful was her visage'
 (255) bha a gnùis brèagha
 'her visage was beautiful'

3.1.3 Locatives and Possessive Predication

Generic locatives are formed similarly to predicative adjectives: *THA* + NP^{SUB} + LOC. The location can take the form of a prepositional phrase or demonstrative:

- (256) bha sinn shìos aig a' chladach
 be-PAST 1P down-LOC at ART shore
 'we were down at the shore'

- (257) tha an càr a-muigh
 be-PRES ART car out-LOC
 'the car is outside'

There is no verb 'to have' in ScG. Similarly to Russian and Estonian, possessives are formed as locatives using one of the prepositions marking location, usually *AIG*:

- (258) tha leabhar agam
 be-PRES book at-1S
 'I have a book'

- (259) tha an leabhar aig Mairead
 be-PRES ART book at Margaret
 'Margaret has the book'

In general, if the locative-phrase NP is higher on a scale of animacy than the subject NP, the relation is one of possession. Otherwise, it is understood as location (Macaulay 1992: 182). If they are equally animate, either reading can obtain:

- (260) bidh Dòmhnall aig Sìle 'Donald will be at Sheila's/Sheila will have Donald'

3.1.4 Existentials

There is no presentative construction in Gaelic as such. Deixis determines whether a clause is read as introducing a participant into discourse as can be seen in the difference between the translation of the following two clauses:

- (261) tha eun aig an uinneig
 be-PRES bird at ART window-D
 'there is a bird at the window'

- (262) tha an t-eun aig an uinneig 'the bird is at the window'

So-called 'pure existentials', performing a substantive function, are formed using the preposition *ann* 'in it':

- (263) a bheil Dia ann
 QP be-PRES-DEP God EXIST
 'is there a God'

- (264) bidh an sneachd ann 'there will be snow'

- (265) nach bi sibh ann 'won't you be there'

Ann is a ubiquitous word in the language, taking its function largely from context and pragmatics. One of its specialised uses is in emphasising negation, possibly linked with its existential function:

- (266) cha bhi mi ag òl ann
 NEG be-INDEF1-DEP 1S PROG drink-VN EXIST(?)
 'I won't be drinking at all'

- (267) chan eil mi eòlach air an duine ann
 NEG be-PRES-DEP IS knowledgeable on ART man EXIST(?)
 'I don't know the person at all' (*eòlach* adj. 'having knowledge of')

3.2 Constituent Order and Grammatical Relations

It is useful to posit a three-way division between the basic semantico-syntactic roles S, A, and P as follows: 'S' is the single nominal argument of a one-argument clause ('intransitive' clause); 'A' is the most-agentive argument of a many-argument clause ('transitive' clause); and 'P' is the most PATIENT-like argument of a many-argument clause. Scottish Gaelic, like other nominative-accusative languages, groups A and S together morphosyntactically while distinguishing these from P. The pragmatically neutral constituent order is VS/VAP (traditionally VSO) as can be seen in the following:

- (268) V S
 ghluais an cat
 move-PAST ART cat-N
 'the cat moved'
- (269) V A P
 ghluais an cat am ball
 move-PAST ART cat-N ART ball-N
 'the cat moved the ball'

If the initial verb is an auxiliary, the order becomes AUX-S-VN or AUX-A-VN-P:

- (270) AUX S VN
 tha an cat a' gluasad
 be-PRES ART cat-N PROG moving-VN
 'the cat is moving'
- (271) AUX A VN P
 tha an cat a' gluasad a' bhuill
 be-PRES ART cat-N PROG moving-VN ART ball-G
 'the cat is moving the ball'

A morphologically clear division of P from A and S is seen in the last example. In transitive clauses with an auxiliary and verbal noun, the nominal argument P is traditionally marked by genitive case (see discussion in A§2.2.1.4) or, if pronominal, incorporated into the aspectual particle of the verbal noun (see A§2.3.1.1). Although case marking does not always distinguish A from P in Gaelic clauses, syntax or pragmatics always does. Thus it is appropriate to posit the grammatical relations 'subject' and 'object' in Scottish Gaelic, 'subject' being a combination of A and S and 'object' being P alone.

More evidence of nominative/accusative marking can be seen in the way that ScG handles complement argument omission (adaptation of examples in Payne 1997: 162-3):

- (272) a. tha Ceit, airson , pòsadh 'Kate wants to marry'
 S omission
- b. tha Ceit, airson , Pòl a phòsadh 'Kate wants to marry Paul'
 A omission

c. tha Ceit airson Pòl pòsadh ⁶⁰	'Kate wants Paul to marry'	none
d. tha Ceit, airson Pòl a, pòsadh	'Kate wants Paul to marry her'	none

The word *airson* 'for' is a complex preposition that is interpreted in an optative sense when associated with an AGENT. It can take either an NP, or a non-finite verbal complement as in example (272) above. In (a) we see that the S argument of an intransitive complement such as MARRY(x) can be omitted when it is understood to be co-referential with the subject of *airson*. The A of the transitive complement clause MARRY(Kate, Paul) in (b) can be similarly omitted when co-referential with the subject of the matrix clause. In (d) however, this cannot happen since the subject of the complement clause is not co-referential with that of the matrix clause. The object of the complement clause must be made explicit, as it is here by a 3S-FEM possessive pronoun. This is a case of a syntactic process which groups S and A together while treating P separately. Incidentally, the *a* in (b) is not a possessive pronoun but an agreement particle (*a^L*) which indicates that the non-finite verb has a full NP object associated with it. We will return to these examples and non-finite complements generally in A§3.2.2 below.

Indirect objects in ScG, as in ditransitive clauses of transaction, are headed by directional prepositions which govern the dative case and indicate either sources or destinations:

- (273) bheir Brìghde an t-airgead do Chailean
 give-INDEF1 Bridget ART money [O] to Colin [IO]
 'Bridget will give the money to Colin'
- (274) bheir Brìghde an t-airgead bho Chailean
 give-INDEF1 Bridget ART money from Colin
 'Bridget will take the money from Colin'

As can be seen from these examples, while English lexicalises transactional direction on the verb, ScG only does so according to the preposition employed.

3.2.1 Postposition of Weak Pronominal Objects

It is normal for non-contrastive pronominal objects in Scottish Gaelic to occur last in clauses, after adjuncts and prepositional phrases. Consider the following examples:

- (275) a. chunnaic mi an duine sin san eaglais an-diugh
 'I saw that man in church today'
 b. chunnaic mi esan san eaglais an-diugh
 'I saw *him* in the church today'
 (cf. *chunnaic mi san eaglais an-diugh esan)
 c. chunnaic mi san eaglais an-diugh e
 *'I saw in the church today him'

Full object NPs must occur immediately after the subject in clauses where the lexical verb occurs first (i.e. those that are not aspectually marked). Likewise, the same obligatory order must occur with emphatic pronominal objects. Yet, as seen in the last example, non-stressed pronominal objects can and usually do occur in a clause-final position. Adger (1997; see also Ramchand 1997) indicates that this is a prosodically-motivated phenomenon common to the Goidelic languages in general and not, as previous researchers have claimed, syntactically-motivated. The tendency, however, is tempered by the length of intermediating elements:

⁶⁰ For most speakers today, the normal way of saying this would be *tha Ceit airson 's gum pòs Pòl* or *tha Ceit ag iarraidh gum pòs Pòl*, which are 2 clause structures.

- (276) chunnaic mi [e] san eaglais ùir [e] ri taobh na h-aibhneadh [e] as dèidh na
coinneimh aig trì uairean feasgar an-dè [ʔe]
'I saw him in the new church [him] beside the river [him] after the meeting at 3
o'clock in the afternoon yesterday [ʔhim]'

Here it would be more intelligible for the pronoun to occur either immediately after the subject NP or between the place adjuncts and the time adjuncts.

3.2.2 Small Clauses

Small clauses, also called reduced clauses, are syntagms that resemble full clauses in certain ways but lack a finite verb. In Gaelic, some predicating elements, especially finite verbs and verbal nouns, but also some prepositions and aspect markers, are able to take non-finite complements. The simplest form of this is as in (272)a above which, semantically, consists of a reduced verbal core having an implicit argument that is co-referential with the subject of the matrix clause. However, as was seen in the other examples, subtle differences in morphosyntax can result in large differences in interpretation. Intransitive reduced cores, as seen in (272)a,c are not marked by lenition. Their single argument is either unstated or, if different from the subject of the matrix clause, marked as a full NP, as in (272)b, or as a possessive pronoun:

- (277) tha Ceit airson a phòsadh 'Kate wants to marry him'
S PRED O VN

The syntax of a transitive non-finite complement, as seen here, is essentially (object-) subject (-ath)-verbal noun. This structure in this example is sometimes called an 'inverted nominal' in Gaelic grammars due to the object coming before the verb, in contrast to more basic transitive clauses such as (271) above.

There is a potential for vagueness when the subject NP is masculine as the agreement particle mentioned above is homophonous with the 3S-M possessive pronoun:

- (278) tha Pòl airson Ceit a mharbhadh

This clause can be interpreted as either 'Paul wants to kill Kate' or 'Paul wants Kate to kill him'.⁸¹ When discussing marriage, who wants the other 'for to' marry may be inconsequential, but with other verbs this is not necessarily so.

Small clauses are, by definition, unspecified for tense but this does not preclude them from taking aspect. In fact, transitive clauses marked for perfect aspect are characterised by their 'inverted nominal' form:

- (279) tha Ceit air Eòin a phòsadh
be-PRES Kate 'after' John AGR marry-VN
'Kate has married John'

Small clauses in progressive aspect are frequently associated with a type of clause combination known as cosubordination, which in simple terms shares some features with subordination and others with coordination. Although the clause linkage marker is a conjunction, the interpretation is temporal and the semantic bond between the two clauses is tighter than it would be in a case of coordination (see A§3.3.3 for more information about these issues):

⁸¹ This is one of two cases where Gaelic groups A and P. The other is in relative clauses (see A§3.3.2.2).

- (280) chunnaic mi Ceit agus i a' pògadh Phòil
 see-PAST 1S Kate and 3S-FEM PROG kiss-VN Paul-G
 'I saw Kate while she was kissing Paul' [Lit. I saw Kate and she at the kissing of Paul']

3.2.3 'Composite Verbs'

There are dative subject (or 'affective') constructions in ScG where the subject NP is obliquely marked while the 'object', if present, is in the nominative/accusative case. Many of these are modal-type expressions, involving obligation, necessity, ability and so-on while others function in the way that psychological verbs do in other languages:

- (281) is toil leam an duine sin
 COP-PRES like with-1S ART man-N that
 'I like that man'
- (282) b' aithne dha d' athair Seonaidh Bàn
 COP-PAST acquaintance for/to 2S-POSS father-D Johnny-N 'White'
 'Your father knew faired-haired Johnny'
- (283) is beag orm marag
 COP-PRES little on-1S blood-pudding
 'I dislike blood pudding'

The predicate is formed by a copular verb which 'joins' with either a nominal or adjectival element. These are often idiomatic in nature, as in (283) above. These 'composite' verbs can take non-finite complements such as a small clause in (284) and an infinitive construction in (285):

- (284) bu toil leam marag a thoirmeasg 'I would like to ban blood-pudding'
- (285) is beag oirre a bhith 'na cabhaig 'She dislikes to be in a hurry' [Lit. 'in her hurry']

Others generally only take verbal complements:

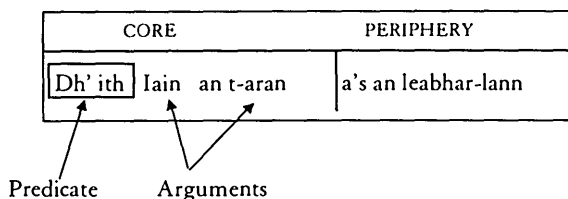
- (286) bu chòir do Sheònaig an càta sin a cheannach
 COP ought for/to Jane-D ART coat that AGR buy-VN
 'Joan ought to buy that coat'

3.3 Clause Combinations

There has been considerable debate on how best to categorise clause combinations in natural language, challenging the traditional definitions of notions such as subordination (Beaman 1984; Foley and Van Valin 1986; Haiman and Thomson 1988; Lakoff 1984; Mathieson and Thomson 1988; Thomson 1984; Van Valin 1985; Van Valin and LaPolla 1997). It has been suggested that part of the problem may rest in an underlying bias in linguistics towards the Indo-European languages which often have overt, lexical realisation of clausal relations. Role and Reference Grammar (RRG: Van Valin and LaPolla 1997 *inter alia*) developed out of an attempt to account for, in the first instance, a wide range of non-IE languages and it has been particularly useful in clarifying this area. The following is a brief introduction to the way that RRG views clause combinations and it will inform the discussion in the next few sub-sections.

In RRG, the clause is construed as having three levels: the **nucleus**, essentially the predicate; the **core**, being the predicate and its arguments; and the **clause**, which is the predicate, its arguments, and associated adjuncts:

The Layered Structure of the Scottish Gaelic Clause



- (287) dh'ith Iain an t-aran a's an leabhar-lann
 cat-PAST Ian ART bread-N in-ART library-N
 'Ian ate the bread in the library'

This example can also be represented as a tree diagram:

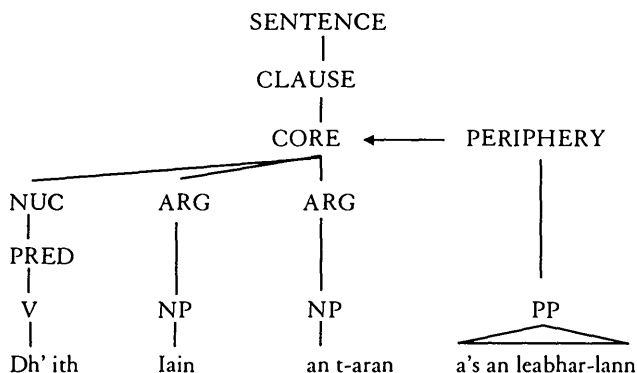


Figure 3: Tree Diagram of the Scottish Gaelic Clause

Juncture is the linking of two or more of these elements: e.g., a nuclear juncture results from the linking of two or more nuclei. Juncture is distinguished from **nexus**, which is the type of relation that obtains between a given juncture. Three different kinds of nexus are posited: **coordination**, **subordination** and **cosubordination**. Theoretically, there are nine different combinations of nexus and juncture, such as nuclear coordination (i.e. nuclear juncture + coordinate nexus), core subordination, clausal cosubordination, and so. As this section is mainly dealing with clause combinations, a discussion of nuclear and core juncture would be digressive. However, the following examples are provided to illustrate, on a general level, what these would be like (for more information about nuclear and core juncture see Van Valin and LaPolla 1997: 442–484):

Nuclear juncture: “single cores containing more than one nucleus ... the multiple nuclei function as a single complex predicate taking a single set of core arguments” (Van Valin and LaPolla 1997: 448):

- (288) thalla is faigh an t-òrd tha an-sin
 away-IMP CONJ get-IMP ART hammer-N be-PRES there
 'go and get me that hammer'

Core juncture: "single clauses containing more than one core. Each core may have its own core argument(s) not shared with the other core(s)" (Van Valin and LaPolla 1997: 448)

- (289) tha mi ag iarraidh am pàipear sin a leughadh
 be-PRES 1S PROG wanting-VN ART paper-N that AGR reading-VN
 'I want to read that paper' [Lit. 'I am at wanting that paper to read']

Clausal juncture: "whole clauses are joined, and each clause may be fully independent of the others" (Van Valin and LaPolla 1997: 448)

- (290) sgioblaich an seòmar agad agus nigh d' aodach cuideachd
 tidy-IMP ART room at-2S CONJ wash-IMP 2S-POSS clothes also
 'tidy your room and wash your clothes too'

One problem with many accounts of clausal relations is that they fail to discriminate between structural and operational dependency. If a clause is structurally dependent, it is unable to 'stand on its own' as an independent utterance. If it is operationally dependent, on the other hand, it is parasitic on the matrix clause for certain grammatical categories such as mood, tense, aspect, or illocutionary force. These are fundamentally different from predicates and their arguments and are referred to as **operators** in RRG. They are grouped according to the specific level of juncture to which they pertain. For instance, aspect operates at the level of the nucleus while tense and illocutionary force operate at the level of the clause, hence they are called nuclear and clausal operators respectively.

Subordinate clauses can manifest themselves as either being embedded (e.g. functioning as a core argument) or modifying the periphery of a matrix clause. In both cases, the defining feature is structural – rather than operational – dependence. Consider the following three utterances:

- (291) cuiridh [eadhon a bhith ag èisteachd ris] dragh orm
 put-FUT even be-INF PROG listen-VN to-3S-M annoyance on me
 'even listening to him annoys me' [Lit. 'even listening to him puts annoyance on me']
- (292) bha dragh orm [nuair a chuala mi a ghuth]
 be-PAST vex on-1S when REL hear-PAST 1S his voice
 'it vexed me when I heard his voice'
- (293) thachair iad rithe [agus i 'dol dhachaigh]
 happen-PAST 3P 'to her' and she PROG-go-VN home
 'they happened upon her while she was going home'

In (291), a non-finite core (*eadhon a bhith ag èisteachd ris* 'even listening to him') is embedded as a subject within a clause. Because it functions as an argument of the predicate (*cuiridh* 'will put'), this is core, not clausal, subordination (although these are often classified together in grammars). Example (292) is a case of clausal subordination: there is a relation between two clauses where an adverbial modifying clause (*nuair a chuala mi a ghuth* 'when I heard his voice') is structurally dependent on and peripherally modifies a matrix clause (*bha dragh orm* 'a vex (bother) was on me'). Example (293) is fundamentally different than the others although Boyle (1973) and Ó Siadhail (1984) would classify it as a type as subordination. The difference is down to the presence of obligatory operator dependence; the non-finite clause relies on the matrix clause for tense and illocutionary force. In RRG terms, this is cosubordinate clausal nexus. This type of non-matrix clause also behaves differently to subordinate ones in certain syntactic tests such as

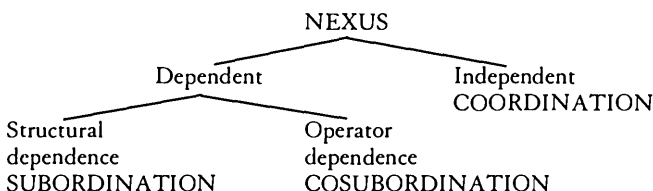
preposing, which is ungrammatical with ScG cosubordinate clauses but possible with subordinate ones:

(294) nuair a bha i 'dol dhachaigh, thachair iad rithe
'when she was going home, they happened upon her'

(295) *agus i 'dol dhachaigh, thachair iad rithe

The concept of finiteness relates primarily to full morphological marking of tense and other clausal operators. In the literature, one frequently encounters discussion of non-finite clauses in relation to verbal nouns, gerunds, participles and such. It seems more accurate to refer to these as cores instead. Clauses will be taken to be syntagms with the structural characteristics mentioned above (A§3.3) for which the status of clausal operators, such as tense and illocutionary, is relevant, even when not specified. The use of 'non-finite clauses' will be limited to instances where a reduced clause is in a relation of operational dependence with a fully finite matrix clause, i.e. cosubordinate clauses. Only they have the requirements for clausal status despite the fact that their clausal operators are parasitically engendered.

To sum up, clausal relations, indeed nexus in general, can be conceived as a three-way division (from Van Valin and LaPolla 1997: 454):



3.3.1 Coordination

There are three main ways of linking constituents of equal syntactic status: conjunction; disjunction; and exclusion. In ScG, these are expressed lexically by a small set of conjunctions. Because many of the conjunctions used in Gaelic are the same for clausal and phrasal constituents, both are included in this section. Conjunctions always occur medially, between the constituents to which they pertain.

3.3.1.1 Conjunction

Conjunction of the type 'a *and* b' is grammaticalised in Gaelic by the word *agus* and its shortened forms '*s*, *a's*, and *is*. Zero marking or juxtaposition is not a frequent strategy in speech although in poetry, adjectives are sometimes strung together one after another. Lists also sometimes show zero marking, but *agus* (or more usually '*s*) is ubiquitous and it is often found tagged on, almost like an affix, any time there is conjunction.

(296) ith do shuipear 's bi sàmhach
eat-IMP 1S-POSS supper and be-IMP quiet
'eat your supper and be quiet'

(297) fhuair mi an aithris agus bheir mi sùil a-nochd oirre
get-PAST 1S ART report CONJ give-INDEF 1S eye tonight on it-F
'I got the report and I will take a look at it tonight'

The word *na* is used with phrasal constituents to convey 'not a nor b':

- #### 3.3.1.2 Disjunction

3.3.1.3 Exclusion

(304) chan eil mi ag iarraidh deoch ach biadh
'I don't want drink but food'

Ach also functions as a 'focus' particle in a type of pseudo-cleft (A§3.4.2.2) and can also indicate concession in constructions such as the following (adapted from Macaulay 1992: 164):

- (305) *thigeadh e ach is e falbh*⁸²
 come-INDEF2 3S-M but 3S-F-EMPH go-VN
 'he would come provided *she* left' [Lit. 'he would come but *she* leaving']

3.3.2 Subordination

3.3.2.1 Complement Clauses

Complement clauses are of a syntactically different status to adverbial clauses. The former serve as logical arguments of a predicate; they can act as 'subjects' or 'objects'. Adverbial clauses, on the other hand, are adjunctive modifiers of a matrix clause, situated in that clause's periphery. In ScG, complement clauses can be headed by verbs, nouns, adjectives and prepositions. They are generally marked by using the dependent verbal complementiser *gu(n/m)/gur* and, if negative, *nach* rather than *cha(n)*. Examine the following:

Verb-headed complement clause

- (306) *chan eil mi a' creidsinn gun tèid e ann*
 NEG be-PRES-DEP 1S PROG believe-VN COMP go-INDEF1 3S-M 'there'
 'I don't believe that he will go there'

In (306), what the subject 'does not believe' is 'that he will go'; the complement clause acts as a core argument, an 'object'.

Noun-headed complement clause

- (307) *chan eil rian nach fhaic thu sin*
 NEG be-PRES-DEP way NEG see-INDEF1-DEP 2S that
 'there's no way that you don't/can't see that'

Here, the predicate is formed by the noun *rian* 'way' which heads a negative complement clause.

Adjective-headed complement clause

- (308) *tha e fìor gun tuirt e seo*
 be-PRES 3S-M true COMP say-PAST-DEP 3S-M this
 'it is true that he said this'

In this example, there is an empty subject – *e*, the third person masculine pronoun. Complement clauses cannot begin statements in Scottish Gaelic the way they can in English; they cannot act as syntactic subjects. In English, it is acceptable, although rare and mostly confined to planned writing, to find utterances such as 'that he said this is true'. It is more common to postpose the complement clause in English using an empty subject, and this is the only option in Gaelic: 'it is true that he said this'.

Preposition-headed complement clause

- (309) *tha iad airson 's gun dèanadh tu an ceart-uair e*
 be-PRES-INDEP 3P for CLM COMP do-INDEF2-DEP 2S right now 3S
 'they want you to do it right now' [Lit. 'they are for that you would do it right now']

⁸² The most common way of expressing this utterance would be with a conditional clause however: *thigeadh e nam falbhadh is* 'he would come if she would leave'.

The complement clause in (309) again functions as an argument of the predicate (*airson* 'for'), specifically its object.

Non-finite complements have a similar distribution but, unlike finite complements, some can also occur in first position, as for instance in (291), repeated here:

- (310) cuiridh [eadhon a bhith ag èisteachd ris] dragh orm
 put-FUT even be-INF PROG listen-VN to-3S-M annoyance on-2S
 'even listening to him annoys me' [Lit. 'even listening to him puts an annoyance on me']

The non-finite complement *eadhon a bhith ag èisteachd ris* functions as the syntactic subject and occurs in the place normally reserved for subjects in the language. Often, however, these are postposed with empty subjects as in the following example:

- (311) chòrdadh e glan rium an càr sin a dhraibheadh
 enjoy-INDEF2 3S-M clean-ADV to-1S ART car that AGR driving-VN
 'I would really enjoy driving that car' [Lit. 'it would be clean enjoyable to me to drive that car']

In (312), the past participle *toilichte* heads a non-finite complement with a pronominal argument. Literally, this could be construed as 'we are pleased your seeing'. The following example is identical except that the predicate is formed by the complex preposition *airson* and there is a temporal adverb modifying the non-finite complement.

- (312) tha sinn toilichte ur faicinn
 be-PRES 1P pleased 2P-POSS seeing-VN
 'we are pleased to see you'

- (313) tha sinn airson ur faicinn a-rithist
 'we want to see you again'

3.3.2.1.1 Indirect Questions

Indirect questions have the same syntax as interrogative clauses (see A§2.3.1.2 Mode):

- (314) a bheil fhios aca càite an deach e
 QP be-PRES-DEP of knowledge at-3P where go-PAST-DEP 3S-M
 'do they know where he went'
- (315) chan eil iad cinnteach an do dh'fhalbh i mar-thà
 NEG be-PRES-DEP 3P sure QP went-PAST-DEP 3S-F already
 'they are not sure if she left already'

3.3.2.2 Adverbial Clauses

Adverbial clauses in Scottish Gaelic cover a wide range of semantic relations. Adverbial clauses are subordinate in the senses mentioned in §3.3 and are generally able to be preposed, clefted, and, to some extent, nested within the constituents of the matrix clause.

The morphosyntax of subordination in Scottish Gaelic is of four different types, relating to the type of verbal complement taken by a particular subordinator:

Type I: subordinator + independent verb form

In Type I, the subordinate clause is only so semantically; if the subordinator were to be removed, the clause could stand on its own. Some of these subordinators have a variant form in Type II or IV, but not necessarily vice versa. There may be functional motivations for this variation, which are explored in the following section.

- (316) tha fhios a'm glè mhath oir bha mi aig ceann an ròpa
 be-PRES knowledge at-1S very well for be-PAST-INDEP 1S at end ART rope-G
 'I know very well for I was at the end of the rope'

Type II: subordinator + agus + dependent verb form

- (317) a chionn 's gur ann aig a' Chumhang a thugadh na buillean
 since and COP-DEP FOC at ART Strait REL give-PAST-IMPERS ART hit-P
 'since it was [Lit. 'is'] at the Strait that the blows were exchanged'

cf.: *a chionn, is ann aig a' Chumhang...* [Type I: independent copula used]

Type III: subordinator + relativiser + relative verb form⁸³

- (318) on a tha thu 'dol a dh'fhalbh
 since REL be-PRES 2S PROG-going to depart-INF
 'since you are going to depart...'

Type IV: subordinator + dependent verb form

- (319) cha d'fhuair e a' chroit seach gu robh sublet aige
 NEG get-PAST-DEP 3S-M ART croft since COMP is-PAST-DEP sublet at-3S-M
 'he didn't get the croft since he had a sublet'

Reason

Scottish Gaelic has several different 'reason' subordinators. Some of these are as follows:

<i>a chionn</i> ⁸⁴	I, II, IV
<i>air sàillibh</i>	I, II, IV
<i>air sgàth</i>	II
<i>airson</i>	I, II, IV
<i>air tàillibh</i>	(variant of <i>air sàillibh</i>)
<i>(bh)o</i>	I
<i>do bhrìgh</i>	IV
<i>leis</i>	IV
<i>oir(eadh)</i>	I
<i>ri linn</i>	II, IV
<i>seach</i>	IV
<i>thoradh</i>	I, IV

⁸³ There is a distinct verbal form used with weak verbs in the 1st Indefinite ('present-future'); otherwise, the verb is identical to the independent form. Consider the following: *ma dh'fhàgas tu e* 'if you (will) leave it'; *ma dh'fhàg thu e* 'if you left it'; c.f. *fàgaidh mi e* 'I will leave it' and *dh'fhàg mi e* 'I left it'.

⁸⁴ Many speakers today use *a chionn* 's itself as a Type I subordinator: e.g. ...*a chionn 's, is ann aig a' Chumhang a thugadh na buillean* (as above).

There is dialectal variation in the types of complement available to each subordinator, and there are certainly fine distinctions in meaning between these words. Type I usage is much closer to coordinate nexus and, as expected, cannot be preposed (or clefted):

(320) rinn mi sin a chionn bha fearg orm
do-PAST IS that because is-PAST angry 'on me'
'I did that because I was angry'

(321) *a chionn bha fearg orm rinn mi sin
'because I was angry, I did that'

If, however, a type II or IV form is used, the utterance is acceptable:

(322) a chionn's gun robh fearg orm rinn mi sin

Part of the distinction between these forms and their acceptability in different positions may involve what Chafe calls 'bound' and 'free' adverbial clauses (Chafe 1984). Bound adverbial clauses are those that occur, within spoken language, in the same intonation unit as the matrix clause and, in writing, within the same punctuation unit. What this means is that, in both speaking and writing, bound adverbial clauses are joined to the main clause in one combined chunk; there is no pause or break between the two clauses.

Free adverbial clauses on the other hand are those that are separated from the matrix clause in speech, by a pause, and in writing, by a punctuation mark (usually a comma). In other words, with a free adverbial clause there is a marked break between it and the matrix clause. The following examples in English are from Chafe (1984: 439):

- | | |
|--------------------------|---|
| (a) Preposed and bound: | Because it has such a big memory I decided to buy it |
| (b) Postposed and bound: | I decided to buy it because it has such a big memory |
| (c) Preposed and free: | Because it has such a big memory, I decided to buy it |
| (d) Postposed and free: | I decided to buy it, because it has such a big memory |

Chafe says that when a free adverbial clause is postposed, it is like an afterthought with its relationship to the matrix clause closer to coordination than subordination. Returning to the Scottish Gaelic examples above, there seems to be a tendency to assign free adverbial clauses to Type I structures and bound adverbial clauses to Types II, III and IV. The more 'coordinate' status of Type I subordinators explains their inability to prepose or cleft. This is a case of iconicity: the closer the semantic link between two clauses, the more integrated is the morphosyntax making the link explicit.

NB: Some of the above subordinators can be used with phrasal constituents to convey the sense 'because of', 'due to' or 'for the sake of':

(323) dh'fhalbh Màiri gu cliobhar air sgàth na sabaid
leave-PAST Mary ADV clever due to ART fighting-G
'Mary left quickly due to the fighting'

Concessive Clauses

These are realised by the Type I subordinator *ged* 'although/even' as in:

(324) [ged a tha sinn sgìth] cumaidh sinn oirnn
'although we are tired, we will keep going' [Lit. '...we will keep on us']

Conditionals

In all but INDEF2 and the copula, simple positive conditionals are initiated by the Type III subordinator *ma* 'if':

- (325) [ma dh'fhàg thu an-shin e] bidh e air a ghoid a-nis
 if leave-PAST 2P there-LOC 3S-M bi-INDEF1 3S-M after 3S-POSS steal-VN now
 'if you left it there it will be stolen (by) now'

The copula takes the form *mas* (*ma* 'if' + (*a*)s - the relative form of the copula) as in the following:

- (326) nì mi e [mas urrainn mi⁸⁵]
 do-INDEF1 1S 3S-M if-COP able 1S
 'I'll do it if I can'.

In hypothetical conditionals, the second indefinite is used in conjunction with the subordinator *nan/nam* (Type IV):

- (327) dhèanainn e [nam b' urrainn mi]
 do-INDEF2-1S 3S if COP-INDEF2 able 1S
 'I would do it if I were able'

A special use of *nan/nam* is with the past definite or second indefinite to convey an optative sense:

- (328) nan robh bàta agam bheirinn sgriob a dh'Uibhist
 if is-PAST-DEP boat at-1S take-INDEF2-1S trip to Uist
 'if (only) I had a boat I would take a trip to Uist'.

In counterfactuals there is an interesting intermingling of the tense, aspect and mode system. The clause of condition begins with the subordinator *nan/nam* and a verb in the 2nd indefinite or the past definite, as in the previous example. The clause of result is in the past definite. Gaelic grammarians referred to this as the 'modal preterit' although it is not to be found in any of the grammars.⁸⁶ In essence, there is a clause expressing an irrealis condition juxtaposed with one expressing a realis result:

- (329) nan robh thu air an doras a dhùnadh
 if be-PAST-DEP 2S PERF ART door AGR closing-VN
 'if (only) you had closed the door'
- cha robh a' chaora air faighinn a-steach
 NEG be-PAST-DEP ART sheep PERF getting-VN in-DIR
 the sheep would not have got in'
 [Lit. if you were after the door closing the sheep was not after getting in]

⁸⁵ Although the 'standard' usage is COP *urrainn dhomh*, etc., in Uist at least, it is possible to use the straight pronouns *mi*, *eli*, *sinn*, and *iad* at least (i.e. I have heard or read these usages).

⁸⁶ My thanks to Professor William Gillies for this information.

- (330) nam bithinnsa air falbh an uair sin bha thusa glacte
 if be-INDEF2-1S-EMPH PERF leaving-VN ART time that is-PAST 2S-EMPH locked
 a-muigh
 out-LOC
 'if I had been away [by] then, you would have been locked out'

The previous example could also be expressed in the following way, but with less vivid impact, perhaps because the result clause is irrealis rather than realis:

- (331) nam b' e 's gun robh mi air falbh an uair
 if COP-INDEF2 3S CLM COMP be-PAST-DEP 1S PERF going-VN ART time
 'if it had been that I had been gone then
 sin bhiodh tusa air do ghlacadh a-muigh
 then be-INDEF2 2S-EMPH PERF 2S-POSS locking-VN out-LOC
 'you would have been locked out'
 [Lit. if it had been that I was after going (at) that time you would have been after your locking out]

Negative conditionals are formed by *mura* (Type IV) corresponding to 'if not' or 'unless':

- (332) mura bi iadsan ann cha tig mi fhèin nas motha
 if not be-INDEF1 3P-EMPH there NEG come-INDEF1-DEP 1S REFL either
 'if it will not be there I myself won't come either'

The two comparatives of size 'bigger' and 'smaller' (A§2.2.4.4) can be used idiomatically as well: *nas motha* can mean 'either' as in (332) and *nas lugha* 'smaller/less' can be used to express 'unless':

- (333) nas lugha na tha e shuas an staidhre
 unless that is-PAST 3S-M LOC-up ART stair
 'unless he is upstairs' [Lit. 'less (lest?) than he is upstairs']

Manner

These clauses usually involve *mar* 'as/as if', which can be Type III or IV with a corresponding difference in meaning:

- (334) bha i ga giùlain [mar a sheall mi dhut]
 is-PAST 3S-F PROG-3S-F-POSS carrying-VN as REL show-PAST 1S 'to you'
 'she was carrying herself the way that I showed you'
 (335) bha i ga giùlain [mar gun robh i bochd]
 is-past 3s-F PROG-3S-F-POSS carrying-VN as COMP be-PAST 3s-F ill
 'she was carrying herself as if she were ill'

Location

The subordinator *far* 'where' is used in dependent structures⁸⁷ such as the following:

- (336) chunnaic mi [far an do chuir thu e]
 see-PAST 1S where put-PAST-DEP 2S 3S-M
 'I saw where you put it'

⁸⁷ Technically, these are headless relative clauses.

However, *càite* 'where' is used in indirect questions, such as with *fios* 'knowing':

- (337) tha fhios a'm [càite an do chuir thu e]
 be-PRES knowledge at-1S where put-PAST-DEP 2s 3s-M
 'I know where you put it'

Temporal Subordinators

The following are the most common, glossed along with their types:

(bh)o(n)	'since'	Type I, IV
aon uair	'once'	Type II
fhad 's	'while'	Type III
gun	'until'	Type IV
gus am/an	'until'	Type IV
mun/mus	'before'	Type IV
nuair	'when'	Type III

Clauses headed by these subordinators are able to be preposed and clefted:

- (338) [mus tàinig Alasdair a-nuas] ghabh i norrag bheag
 before come-PAST-DEP Alasdair down take-PAST 3S-F nap small
 'before Alasdair came over she took a wee nap'

cf. [ghabh i norrag bheag] mus tàinig Alasdair a-nuas

- (339) 's ann [mus tàinig Alasdair a-nuas] a ghabh i norrag bheag
 'it's before Alasdair came over that she took a small nap'

Simultaneity can be expressed using *fhad 's* or a cosubordinate expression with *agus* (A§3.3.3):

- (340) bha i 'na cadal [fhad 's a bha Alasdair a' feitheamh rithe]
 is-PAST 3S-F 'in her' sleep 'while' REL is-PAST Alastair PROG waiting-VN to-3S-F
 'she was asleep the whole time Alastair was waiting for her'

3.3.2.3 *Relative Clauses*

Relative clauses are a case of **noun phrase subordination**; they function as nominal modifiers and do not constitute clausal nexus as such. In other words, relative clauses serve to expand or modify the meaning of noun phrases rather than clauses. This can be seen in an example such as *the man [who stole your parasol] is leaving the train*, where the relative clause (in brackets) modifies the NP 'the man' and, together, they form a constituent. This contrasts to examples of clausal nexus such as *the man left the train [when he saw us]* where the subordinate clause serves to modify the main clause as a whole, not just a single part of it.

In ScG, relative clauses always occur post-nominally and are headed by the relativiser *a*:

- | | |
|------|--------------------|
| HEAD | RESTRICTIVE CLAUSE |
|------|--------------------|
- (341) an trusdar a [ghoid ø an làraidh aig Aonghas Iain]
 ART vagrant REL stole NP_{REL} ART lorry at Angus John
 'the vagrant who stole Angus John's lorry'

As can be seen, the case recovery strategy is to simply leave a gap (\emptyset) in the restrictive clause. Because of ScG's VSO order, this can result in problems recovering grammatical relations (A and P) if the restrictive clause is headed by a transitive verb. Of course, discourse pragmatics almost always disambiguate such cases:

- (342) sin an duine a chunnaic mi
 that-(COP) ART man REL see-PAST 1S
 'that's the man who saw me' or 'that's the man whom I saw'

All positions on Keenan and Comrie's relativisation hierarchy can be relativised using the relativiser-gap strategy except possessors, which use a resumptive possessive pronoun. In obliques, the tendency is towards pied-piping + \emptyset , but the relativiser-gap strategy is sometimes used in coordination with a resumptive prepositional pronoun:

Table 139: Relativisation strategies

EXAMPLE	GLOSS	POSITION	STRATEGY
<i>sin an gille a tha bochd</i>	that's the boy who is ill	subject	REL + \emptyset
<i>sin an gille a chunnaic Ceit</i>	that's the boy who Kate saw (also 'that's the boy who saw Kate')	object	REL + \emptyset
<i>sin an gille a thug an litir do Chèit</i>	that's the boy who gave the letter to Kate	indirect object	REL + \emptyset
<i>sin an gille air an do shuidh Ceit</i>	that's the boy on whom Kate sat	oblique	PREP + \emptyset
<i>sin an gille a shuidh Ceit air</i>	that's the boy who Kate sat on	oblique	REL + \emptyset + PREPPN _{RES}
<i>sin an gille a tha a mhàthair bochd</i>	that's the boy whose mother is ill [Lit. 'that's the boy who his mother is ill']	possessor	REL+ POSSPN _{RES}

3.3.3 Cosubordination

There are constructions in ScG that have more than a passing resemblance to both coordination and subordination. They share with the former the overt expression of a conjunction (*agus*) although semantically, they can resemble a relative clause or an adverbial adjunct. Compare the following, in which each first example is a case of cosubordination and each second one, either subordination or coordination.

'Participle Type'

- (343) dh'fhalbh Alasdair 's an t-acras a' tighinn air
 leave-PAST Alastair CONJ ART anger PROG coming-VN 'on him'
 'Alasdair left with hunger coming on him'
- (344) dh'fhalbh Alasdair 's bha an t-acras a' tighinn air
 'Alasdair left and hunger was coming on him'

'Reason Type'

- (345) dh'fhalbh Alasdair agus i 'na suain
 leave-PAST Alastair CONJ she in-3S-F slumber
 'Alasdair left because she was fast asleep' [Lit. '...and she in her slumber']
- (346) dh'fhalbh e a chionn 's gun robh i 'na suain
 'he left because she was fast asleep'

'Relative Clause Type'

- (347) bha duine ann 's Aogh mar ainm dha
 is-PAST man EXIST and Hugh as name 'for him'
 'there was a man who was called Hugh' [Lit. '...and Hugh as a name for him']
- (348) bha duine ann agus bha Aogh mar ainm dha
 'there was a man and Aogh was his name'

'Temporal Type'

- (349) chaidh a chall a's a' chèò 's e ag iasgach aig an àm
 go-PAST 3s-M-POSS lose-VN in the mist-D and he PROG fishing at ART time
 'he got lost in the mist when he was fishing' [Lit. '...and he at fishing at the time']
- (350) chaidh a chall a's a' chèò agus bha e ag iasgach aig an àm
 'he got lost in the mist and he was fishing at the time'

The problem with analysing the first clause in each case as coordinate is that they are non-finite, yet like coordinate clauses they cannot be clefted or fronted. They cannot be regarded as subordinate because of these facts in addition to other syntactic evidence. Structurally, they are reduced clauses with *agus* in utterance-initial position - the place usually reserved for the finite verb. They neatly fit the definition of cosubordination in RRG as they are operationally dependent on a matrix clause for the categories of tense and illocutionary force (which are 'clausal operators'). It is not appropriate to call *agus* a subordinator, as Boyle (1973) and O Siadhail (1984) do. Formally, it is simply a conjunction, which has the effect of being "a virtual red flag, signalling to the hearer, 'these ideas are related somehow, guess how'" (Lackoff 1984: 487). Because an explicit relation is not specified, only that there *is* a relation, a wide variety of semantic functions can be, and have been, ascribed to constructions of this type. However, the semantic interpretation of any particular case is simply a product of human logic which, given the pairing of two propositions, tries to make sense of the association. So, what we have is not logical conjunction, but clausal cosubordination. This is a clear example of the need to keep formal and functional categories distinct.

3.4 Discourse Phenomena

There is hardly any mention in the literature of the way that Scottish Gaelic treats the morphosyntactic coding of discourse phenomena. It is certain that the researcher who sets out in earnest to investigate this area will be amply rewarded by new insights into the language. As this grammar is a sketch, and this is a complex area, it is only possible to give it cursory treatment here.

3.4.1 Referential Devices

There is no specific construction as such for introducing new referents into discourse equivalent to *there* + *COP* in English, but indefinite deixis typically indicates that the referent of an NP is unidentifiable. Normally, brand-new referents are anchored in some way, such as to a possessive pronoun in the following construction:

- (351) tha piseag ùr againn
 be-PRES kitten new at-1P
 'we have a new kitten'

If there is not some other kind of predicate present for anchoring purposes of NP referents that are unidentifiable, *ann* occurs as a sort of generic 'existence' particle

(A§3.1.4). This is the common fare of 'world creation' genres such as storytelling, also used to state that something simply exists (see A§3.1.4):

- (352) bha fuamhaire ann is ceithir làmhnan air
 is-PAST giant EXIST CONJ four arms on-3S-M
 'there was a giant with four arms' [Lit. 'there was a giant and four arms on him']

There is a construction, particularly encountered in narrative, for introducing referents which are to figure significantly in following discourse (taken from text in §4):

- (353) ò bha an t-each mòr geal a bha an-seo a' tighinn aca
 oh is-PAST ART horse-N big white REL be-PAST here PROG come-VN at-3P
 'oh here was this big white horse coming after them'

For referents which are new to discourse but accessible, a left-detachment construction may be used, especially in speech. This often occurs with the phrase *a bheil fhios agad* 'do you know...', prompting the addressee to check for an 'old file' dealing with the information to follow:

- (354) [a bheil] fhios a'd an taigh mòr a tha siud
 QP be-DEP knowledge at-2S ART house big REL be-PRES there
 'you know that big house (that is) there

san Àird chaidh a reic an-dè
 in-ART 'Aird' go-PAST 3S-POSS selling-VN yesterday
 in Aird it got sold yesterday'

Active, identifiable referents are usually coded as pronominals with referent switching (and contrast in general) indicated by emphatic affixes:

- (355) agus cha robh na daoine, a' tuigsinn o'n t-saoghal
 and NEG is-PAST-DEP ART men PROG understand-VN from-ART world
 'and the men didn't understand what in the world

gu dè an aon rud a bha na boireannaich, a' faighinn
 what ART one thing REL is-PAST-DEP ART women PROG get-VN
 (was) the one thing that the women were getting

a bharrachd air a' bheagan a b' urrainn dhaibh-san,
 more on ART little REL COP-INDEF2 able 'to them'-EMPH
 on top of the little that *they* were able

a thoirt dhaibh,
 AGR give-VN 'to them'
 to give to them'

Here, there are two referents (*na daoine* 'the men'; *na boireannaich* 'the women') and two anaphors – prepositional pronouns – referring to them (*dhaibh-san* 'to them [emphatic]'; *dhaibh* 'to them'). There is no differentiation for gender in the third person pronoun and in order to make it explicit that the first occurrence of the anaphor refers to *na daoine* rather than the immediately preceding referent *na boireannaich*, the emphatic suffix *-san* is used. Functionally, in this kind of example, it seems to operate as a form of switch reference. Continuing reference is [-emphatic] while switch reference is [+emphatic].

3.4.2 Highlighting of Constituents

'Highlighting' will be understood here as the means by which a language enables certain constituents to stand out from others in a clause, sentence, or piece of discourse.⁸⁸ In Scottish Gaelic this is accomplished primarily by word order and/or an emphatic suffix.

3.4.2.1 The Emphatic Suffix *-sa/-se*

Dorian (1999) says that this suffix is used to 'highlight contrasts, to place emphasis, and generally to mark the speaker's focus [as well as indicating] changes of focus as the speaker takes conversational turns or moves along in a narrative' (p. 7). In some dialects, it is also used as, or synonymously with, the proximal deictic *seo* (A§2.2.2.2):

- | | | | |
|-------|------------|------|----------------|
| (356) | am fear | seo | am fear-sa |
| | ART man-N | this | ART man-N-EMPH |
| | 'this one' | | 'this one' |

But most commonly, it is used to highlight and track pronominal constituents, as in (355) above. Intonation does not seem to be implicated in pronominal emphasis, making ScG very different from English in this regard. Interestingly, as Dorian (1999) demonstrates, it is possible to have a long stretch of narrative with frequent use of emphatic suffixes for the purposes of maintaining intensity and interest. In comparison, a comparable amount of suprasegmental emphasis in English would be disruptive for the listener.⁸⁹ One of the most frequent uses of this is with the defective verb *arsa*⁹⁰ 'quoth', which always takes the emphatic form of a pronominal subject: e.g. *arsa mise; ars' esan*.

Two examples:

- (357) 's e leabhar gu math doirbh a tha an-sin Iain tha mise a' smaointinn sin
 it's a book very difficult that is there Ian is I at thinking that
 co-dhiubh
 anyway
 'that's a very difficult book John I think that anyway'
- (358) a bheil na caoraich aig Aonghas an lùib nan caoraich [sic] agadsa
 are the sheep at Angus in fold the sheep at you
 'are Angus' sheep together with *your* sheep?'

It is difficult to translate the exact sense of the suffix as the emphasis is often of a different quality and not always as pronounced as intonational emphasis is in English. For instance, in (358) above, the gloss seems to indicate an expression of disbelief on the part of the speaker, but this is not the case. Rather, because of the presentative nature of the utterance, the emphatic suffix is needed to activate the referent coded by *agadsa*, clearly identifiable but previously inactive.

3.4.2.2 Highlighting and Word Order

Syntactic highlighting in Scottish Gaelic is represented by clefts of various forms and the occurrence of constituents in non-canonical (preposed and postposed) position.

⁸⁸ See Miller (2000) for a typology of focus in the languages of Europe, including ScG examples—supplied by the current author—and Irish.

⁸⁹ It is also possible in writing to indicate this kind of suffix based emphasis, without having to resort to *italics* as in English.

⁹⁰ *Arsa* occurs so frequently in narrative, fragmenting clauses into 'bite-sized chunks', that it must serve to aid comprehension on the part of the addressee.

Clefts

There are four basic kinds of clefting constructions in Gaelic, all employing the copula: Nominal-type Clefts (*is e*); Non-nominal Clefts (*is ann*); Deictic Clefts; and WH- clefts. Here we will make reference to the structure of each type of cleft, the types of constituents that can serve as cleft heads, and some mention each cleft-type's basic discourse functions.

1) *Nominal-type clefts* serve to give prominence to NPs and nominalised elements. Like the others, they are bi-clausal and have the form COP + 3S-M + CLEFT HEAD [REL + CLEFT CLAUSE] as can be seen in the following:

- | | | | |
|-------|----------|---------------------------------------|---------------------|
| | | CLEFT HEAD | CLEFT CLAUSE |
| (359) | is e | Uibhisteach | [a chaidh ann] |
| | COP 3S-M | Uist person | REL go-PAST 'in it' |
| | | 'it was a Uist person who went there' | |

- | | | | |
|-------|---|----------------------|---------------------------------------|
| | CLEFT HEAD | | CLEFT CLAUSE |
| (360) | is e | gun do dh'ith | mi am biadh sin [a thug orm a bhith |
| | COP-3S-M | CMPL eat-PAST-DEP 1S | ART food that REL give 'on me' be-INF |
| | cho bochd] | | |
| | so ill | | |
| | 'it is that I ate that food which caused mi to be so ill' | | |

This type of cleft can closely resemble the proper inclusion/identificational construction, (covered in A§3.1.1) when the cleft clause contains *ann* 'there/in it' as in (359) above, but it should be distinguished from it. This is easily accomplished as most nominal-type clefts have a related, non-marked allosentence whereas most identificational constructions do not:

Identificational Construction:

- | | | |
|-------|------------------------------------|---|
| (361) | 's e Uibhisteach a th' ann | 'he is a Uist man' [Lit. 'it is a Uist man that is in him'] |
| | *tha Uibhisteach ann ⁹¹ | 'a Uist man is in him' |

Cleft Construction

- | | | |
|-------|-------------------------------|-----------------------------------|
| (362) | 's e Uibhisteach a chaidh ann | 'it is a Uist man who went there' |
| | chaidh Uibhisteach ann | 'a Uist-man went there' |

One of the functions of this kind of cleft is in marking contrastive focus. In the dialogue below, the negative copula *chan e* serves to contest the proposition that James left with the car and speaker 2 asserts, with 'marked narrow focus' (Lambrecht 1994), that it was actually *Ian* who left with it:

- | | | |
|-------|----|---|
| (363) | 1> | mhothaich mi gun do dh'fhalbh Seumas leis a' chàr |
| | | 'I noticed that James left with the car' |
| | 2> | chan e 's e Iain a dh'fhalbh leis |
| | | NEG-COP-3S-M COP-3S-M Ian REL go-PAST with-3S-M |
| | | 'it isn't ('no') - it is Ian who left with it' |

Is e (or *'s e*) can never be used with non-nominal elements and should not be confused with the copular constructions in A§3.1.2. For instance, the following is ungrammatical:

⁹¹ This is acceptable if the reading is 'a Uist-man is there'.

- (364) *'s e brèagha a bha am boireannach
 COP beautiful REL be-PAST ART woman-N
 'it is lovely that the women was'

In order to front adjectives and other non-nominals, the construction covered next must be used.

2) *Non-nominal type clefts* serve to highlight constituents such as PPs, adverbials, adjectives, and adjuncts but never nominals. They have the form COP + *ann* (+ CLEFT HEAD) [REL + CLEFT CLAUSE]:

- CLEFT HEAD CLEFT CLAUSE
 (365) is ann [an-dè] [a dh'fhalbh Iain leis a' chàr]
 'it is yesterday that Ian left with the car'

This construction can be used to assert contrastive focus as with nominal-type clefts. It can be also used to highlight whole clauses, with a kind of marked sentence focus difficult to render in English:

- (366) 's ann a bha e a' toirt bhuaith
 COP FOC REL is-PAST 3S-M PROG take-VN from-3S-M
 a' chuid a b' fheàrr dhen bhith-beò chaol a bh'
 ART part REL COP-INDEF2 good-COMP of-ART livelihood thin REL is-PAST
 aige
 at-3S-M
 ?'it was that he was taking from him the greater part of the meagre livelihood
 that he had'

This kind of construction serves to sum-up and highlight. The speaker is not simply saying that the ACTOR mentioned took the other man's livelihood, but that *what he did was take from him his livelihood*. However, the construction is intrinsically different to the WH-cleft used here, and much more facile than the translation provided above would indicate.

Between the two cleft types covered up to this point, almost any constituent can be given prominence. Using the simple clause 'Mary will give the money to Neil tomorrow' we can see how this plays out (the following examples adapted from Macaulay 1993: 189–90):

- (367) bheir Màiri an t-airgead do Niall a-màireach
 give-INDEF1 Mary ART money to Neil tomorrow
 'Mary will give the money to Neil tomorrow'
- a. is e Màiri [a bheir an t-airgead do Niall a-màireach]
 'it is Mary who will give the money to Neil tomorrow' (subject NP)
- b. is e an t-airgead [a bheir Màiri do Niall a-màireach]
 'it is the money that Mary will give ...' (object NP)
- c. is e Niall [dhan toir Màiri an t-airgead a-màireach]
 COP-3S-M Neil to-ART give-INDEF1 Mary ART money tomorrow
 'it is Neil to whom Mary will give the money tomorrow' (NP of PP)

- d. 's ann do Niall a bheir Màiri an t-airgead a-màireach
'it is to Neil that Mary will give the money tomorrow' (PP)
- e. 's ann a-màireach a bheir Màiri an t-airgead do Niall
'it is tomorrow that Mary will give the money to Niall' (adjunct)

Verbal cores can also be given prominence in this way:

- (368) 's ann a' toirt an airgid do Niall a bhios Màiri
COP-FOC PROG give-VN ART money-G to Neil REL be-FUT-REL Mary
a-màireach
tomorrow
'it is giving the money to Neil that Mary will be tomorrow'

Or even full clauses:

- (369) 's ann a bhitheas Màiri a' toirt an airgid do Niall a-màireach
'it is that Mary will be giving the money to Neil tomorrow'

3) *Deictic clefts* are essentially equative copular constructions with a demonstrative in subject position followed by the predicate and a relative clause: (COP + 1S-M +) DEM^{SUB} + NP^{PRED} + [REL + CLEFT CLAUSE]. Some examples:

- (370) sin am boireannach a bheir an t-airgead do Niall
that ART woman REL give-INDEFI ART money to Neil
'that's the woman who will give the money to Neil'
- (371) 's e siud an seòrsa rud a tha thu ag iarraidh
COP-3S-M that ART kind (of) thing REL be-PRES 2S PROG want-VN
'that's the kind of thing that you want'

This construction seems to grammaticalise a function similar to reverse WH clefts in English, which either direct the addressee to some earlier point of discourse – as for summing up – or highlight some item from the physical context. It is also possible to have simple deictic equative constructions without a relative clause:

- (372) seo am bràthair aige
this-(COP) ART brother 'at him'
'this (is) his brother'

4) *WH clefts* (sometimes called *pseudo-clefts*) present a reversal of the focus position⁹² inherent to the other cleft types above: in nominal-type clefts it occurs immediately after *e*; in non-nominal type clefts it follows *ann*; in deictic clefts, it occurs after the demonstrative. WH clefts, on the other hand, prepose 'old' or 'predictable' material while postposing the focus.

- PRESUPPOSED —● FOCUS —●
- (373) is e a th' ann ach brògan Choinnich
COP-3S-M REL be-PRES 'in it' but shoes Kenneth-G
'what it is is Kenneth's shoes'

⁹² The 'focus position' serves to mark the unpredictable part of the assertion and is not retrievable from context or previous knowledge (Lambrecht 1994).

In this example, the presupposition is that 'there is something "in it" (there)'. The bit that is added to make a new assertion is that the 'something that is in it' is Kenneth's shoes. The word *ach* is optional – and subject to dialectal variation⁹³ – but serves to explicitly set off the 'old/presupposed' from the 'new/unpredictable' material. The use of *ach* in this way is parallel to the exhaustive listing construction:

- (374) cha robh ann ach brògan Choinnich
 NEG is-PAST-DEP 'in it' but shoes Kenneth-G
 'it was only Kenneth's shoes' [Lit. 'it wasn't but Kenneth's shoes']

Semantically, but not pragmatically, equivalent WH-cleft allosentences can obtain depending on where the copula is placed:

- (375) an aon rud a bhiodh math aice 's e meter fhaighinn
 ART one thing REL be-INDEF2 good at-3S-F COP-3S-M meter get-VN
 'the one thing that would be good for her (to have) it's getting a meter'
- (376) 's e an aon rud a bhiodh math aice (ach) meter fhaighinn
 'it's the one thing that would be good for her (to have) - getting a meter'

If this is revamped as a nominal-type cleft, it is easy to perceive the difference in focus:

- (377) 's e meter fhaighinn an aon rud a bhiodh math aice
 'getting a meter is the one thing that would be good for her'

Non-Canonical Word Order

Preposed nominals are either pre-core ('topicalised'/'fronted') or extracausal, being in a left-detached position (also called 'left-dislocation' or 'NC-Clause': Miller and Weinert 1998):

- | | |
|------------------|---|
| (378) Canonical: | cha ghabhadh Seumas am biadh
NEG take-INDEF2 James ART food
'James wouldn't eat the food' |
| Pre-core: | am biadh _i cha ghabhadh Seumas ø _i
'the food James wouldn't eat' |
| Left-detachment: | am biadh _i cha ghabhadh Seumas e _i
'the food James wouldn't eat it' |

Lambrecht (1994) explores the differences between these in some depth and Van Valin and LaPolla (1997) postulate universal syntactic positions to accommodate the fact that constructions of these types are common in the world's languages. One of the functions of left-detachment is to separate role and reference, which are normally kept apart in speech (Chafe 1992). It can also be used to introduce new referents which, in canonical position, would be too heavy or complex (Prince 1997: 124). Lambrecht (1994) also states that left-detachment can be used to shift between 2 or 3 active referents in discourse and, in this way, can be considered to have a contrastive function.

The phrase *a thaobh* 'as for' is sometimes used in writing where simple left-detachment is felt to be inappropriate for the register at hand, as in English:

⁹³ However, the same cleft type occurs in Irish, also using *ach* as a kind of focus marker. See Ó Siadhail 1989, 336.

- (379) a thaobh an òrain, a rinneadh an 1702 chan urrainn
 as for ART song-G REL make-PAST-IMPERS in 1702 NEG-COP able
 nach e
 NEG-COP-DEP it
 'as for the song that was made in 1702 it can't be that it wasn't...'

Right- and centre-detachment also occur in ScG, especially in unplanned, spoken discourse:

- (380) Right-detachment: cha ghabhadh Seumas e, am biadh,
 'James wouldn't eat it the food'
 Centre-detachment: cha ghabhadh e, Seumas, am biadh
 'he James wouldn't eat the food'

At a basic level, these seem to work to further specify a referent that, although active, might not be immediately identifiable from context if presented only as an anaphor. Gillies (1993: 208) interprets right-detachment in ScG as due to the decaying of the copula or the disappearance of synthetic verbal forms. Evidence in Watson (1927) points to another interpretation: '[the] use of the pronoun with forward reference to a noun which is added as a sort of afterthought or explanation is an old and characteristic Gaelic idiom' (p. 318). He substantiates this with an example from Middle Irish: often an infixed third person singular pronoun is followed by its reference, a full NP, as in *ro-s-foghlaim na gnímu* 'he has learned them the deeds'. In light of this, it seems better to construe right-detachment in ScG as the manifestation of an information structure strategy available to human language rather than a decaying of some sort.

3.4.3 Discourse Particles, Parenthetical Statements, and Interjections

Schiffrin (1987: 37) defines discourse particles as 'sequentially dependent elements which bracket units of talk'. They are characterised as: having low syntactic relevance (they can be deleted from an utterance without impinging on its well-formedness); frequently occurring at the beginning of an utterance; having a range of prosodic contours; and being able to have a variety of meanings depending on discourse characteristics. Scottish Gaelic has several regularly occurring lexical items fitting this criteria. These deserve a proper study in context and it would be premature to comment on their specific functions here.

Discourse particles that have been derived from English are frequent and obvious, especially to the non Gaelic-speaker: *ò* 'oh'; *uell* 'well'; *so* 'so'; *okay*. These seem to have been imported wholesale into the language along with the conventional functions attached to them. Other discourse particles are clearly Goidelic in origin: *seadh* 'right'; *ma-tha* 'then'; *och*; *a-nis(d)* 'now'; *an-dà* 'well'⁹⁴; *co-dhiubh* 'anyway'. These feature prominently in almost any spoken register (except *an-dà*), especially those that are interactional. It would be interesting to see if there are any robust collocations between these and other specific phrases.

There are many fixed clausal expressions in Gaelic which can be used 'parenthetically' to modify the force of a proposition (also see the hedges and emphatics in A§2.2.5.4). Here is a sampling, some of which clearly have an epistemic function:

cha chreid mi 'I believe'⁹⁵
channainnsa 'I (EMPH) would say'

⁹⁴ Uncommon in modern speech, but often seen in folk tale transcriptions.

⁹⁵ This comes from the double negative—*cha chreid mi* + a negative complement clause—but the complement clause is elided.

feumaidh 'it must be'
fhios agad 'you know'
is dòcha 'perhaps'
tha cuimhne agam 'I remember'
tha mi 'creidsinn 'I believe'
tha mi cinnteach 'I am sure'

Gaelic interjections, exclamations and imprecations, caricatured in 'Highland literature' such as *Whisky Galore*, are numerous and colourful. One of the better collections of these is Campbell (1996). Some interjections are seemingly ideophonic, consisting of mono- or di-syllabic words which can reduplicate (and usually do) for greater effect. One remarkable Gaelic characteristic is ingressive speech. This is particularly pronounced, and most frequently heard with interjections, particularly 'aye', which can sound like the speaker, whilst agreeing with his or her addressee(s), is simultaneously inhaling. But multi-syllabic words, phrases, and even entire monologues – in particularly skilled 'circular breathers' – can be performed ingressively. Although this tendency is easily lampooned in imitations of the 'Highlanders', it is not restricted to the Scottish *Gàidhealtachd*. The first time the author heard ingressive affirmation was in Cape Breton. He then noticed it throughout Maritime Canada (settled by Gaelic-speakers in large numbers) and in Ireland. If it were not also a feature of some Scandinavian speech (possibly a indication of its origins), it could be appropriately called the 'Gaelic Gasp'. It is one of the special characteristics of Gaelic orality, with as yet unascertained discourse functions.

Here are some of the more common interjections to be heard in Uist:

<i>àidh àidh</i>	'aye aye' [can be either 'right/yes' or 'yeah right/as if']
<i>isd</i>	'be quiet' [from <i>èisd</i> 'listen']
<i>obh obh</i>	'oh no'; expression of mild disapproval
<i>och och</i>	pity, concern
<i>thalla</i>	'away with you'; disbelief
<i>ud ud</i>	'tut tut'
<i>uell uell</i>	'well well' (should be performed ingressively for full effect)

Exclamations are usually in the form of short phrases, often vocatives which call up a religious personage or make reference to an abstract feeling or concept:

(a) <i>Dhia Dhia</i>	'oh God oh God'
(a) <i>Dhia nan gràsan</i>	'oh God of the graces'
(a) <i>Mhoire Mhàthair</i>	'oh Mother Mary' (S. Uist)
(a) <i>Mhoire Mhoire</i>	'oh Mary Mary' (S. Uist)
(a) <i>Shìorraidh</i>	'oh eternity'
(a) <i>Thì nam buadh</i>	'oh Lord of the virtues'
(a) <i>thiacaïs fhèin</i>	'for goodness sake'
(a) <i>Thighearna 's a Dhia</i>	'oh Lord and God'
<i>an diabhal mise</i>	'the devil me' ['damn']
<i>Dia dham bheannachadh</i>	'God bless me'
<i>Dia dham chuideachadh</i>	'God help me'
<i>Dia dham shàbhaladh</i>	'God save me'
<i>m'eadail</i>	'my dear'
<i>mas fhìor</i>	'yeah right; as if'
<i>ma thogair</i>	'it doesn't matter; who cares'
<i>mo chreach</i>	'my destruction'
<i>mo nàire</i>	'my shame'

It is often said that Gaelic has no obscene curse words. However, the crux of the difference between Gaelic and English, in this respect, may be that lexical obscenity is

relative in Gaelic whereas it tends to be categorical in English. Certainly, part of this has to do with prohibitions concerning the written language, which are not very relevant for a primarily oral language such as Gaelic. A taboo word like 'fuck' is semantically empty (except when having explicit sexual reference); it is offensive because it is understood as being so. In Gaelic, there seem to be few such taboo words. As in many other cultures, children are scolded for blasphemy (many of the above expressions would secure a slap) and for mentioning taboo body parts, but words like *cac* 'excrement' or *galla* 'bitch' do not seem to be vulgar by nature, only by application. Imprecations are more semantically rich and perhaps more varied than they are in English. However, like other spoken genres, it seems that Gaelic imprecation is a dying art and that English curses are quickly supplanting ones like the following (most from MacPherson 1945):

<i>droch bhàs ort</i>	'a bad death on you'
<i>gonadh ort</i>	'wounding to you'
<i>guma h-anmoch dhut</i>	'may it be late ⁹⁶ for you'
<i>gun bhuaidh ort</i>	'without success to you'
<i>mhic an deamhain</i>	'oh son of the demon'
<i>mhic an fhir nach can mi</i>	'son of the one I won't say'
<i>mo mhallachd ort</i>	'my cursing on you'
<i>Taigh Iain Ghròta ort</i>	'House of John O' Groat on you'
<i>taigh na galla(dh) ort</i>	'house of the bitch on you'
<i>thalla is cac</i>	'away and shite' (often used in jest)
<i>thoir an droch àite ort</i>	'take the bad place on you' ['go to hell']

3.5 Influence of English

A sketch of the Gaelic language could hardly be said to be complete without mentioning English. The influence of English on Gaelic has been pervasive and deep, with consequences to every linguistic front from phonetics to discourse and register. We can only touch on some of the general morphosyntactic trends here and much more research is needed on this area, but interested readers should consult Gillies (1980) and Macaulay (1982b; 1986) for more information.

Generally, effects are most pronounced, as expected, in those speakers most immersed in English mass culture. There is certainly a great deal of idiolectal variance in this area with some speakers exhibiting a firm division between Gaelic and English production and others constantly mixing and flowing between the two. Most speakers borrow freely from the English lexicon and use calqued expressions side-by-side native idioms – a natural state of affairs in this type of language contact situation. Whereas in late 19th century or early 20th century Uist, one who spoke mostly English would have been thought to be 'putting on airs', nowadays those Gaelic speakers who do not switch and borrow are similarly marked.

Borrowed lexical items can be either assimilated to Gaelic phonology to various degrees, presented unassimilated, or morphologically calqued (rare and generally contrived for specific purposes):

- Assimilated: *buntàta* 'potato'
- Partly Assimilated: *a' train-adh air a shon* 'training for it' ['train' + *-adh* most productive of the verbal noun suffixes]
- Unassimilated: *na sandwiches agad* 'your sandwiches'
- Calqued: *crios-sàbhalaigh* 'safety-belt' [Lit. *crios* 'belt' + *sàbhalaigh* 'saving' (genitive case)]

⁹⁶ Refers to the belief that dead spirits, fairies, and generally bad things were afoot at night and could catch one unawares.

Grammatical and high-frequency words such as prepositions are rarely borrowed, but almost any content word (e.g. nouns, adjectives, adverbs, and verbal nouns) can be adopted without alteration. Nouns and adjectives may or may not inflect for case depending generally on the speaker/writer's perceived formality of the situation and whether or not the word form is capable of such mutation (see A§2.2.1.4). Calqued phrases are also ubiquitous, with some being particularly tenacious and often preferred over native idioms:

Calque	English Origin	Gaelic Alternative	Gloss
<i>a' dèanamh suas</i>	making up (creating)	<i>a' cruthachadh</i>	'shaping'
<i>a' dèanamh suas d'inntinn</i>	making up your mind	<i>a' cur romhad</i>	'putting before you'
<i>a' gabhail àite</i>	taking place	<i>a' tachairt</i>	'happening'
<i>ga iùsaigeadh</i>	using it	<i>ga chur gu feum</i>	'putting it to use'
<i>gabh thairis</i>	take over	<i>gabh os làimh</i>	'take in hand'
supposed <i>a bhith</i>	supposed to be	<i>tha dùil gum bi</i> <i>tha còir gum bi</i>	'it is expected to be'/'it should be'
<i>tha mi a' coimhead</i> <i>air adhart ris</i>	I'm looking forward to it	<i>tha fiùghair agam ris</i> <i>tha fadachd orm ris</i>	'I have hope of it'
<i>a' coimhead airson</i>	looking for	<i>a' sireadh</i>	searching

Although language-internal processes have levelled many of ScG's indigenous morphological contrasts over the past centuries, the move towards simplification of the verbal and case systems has no doubt intensified through contact with English. With each successive generation of English-speaking bilinguals, a greater amount of these kinds of effects have occurred. Generally, there is a trend towards the preference of periphrasis over inflection. Some examples of these trends:

- the decay of person-marking or impersonal inflection on the verb (A§2.1.1)
- levelling of case and gender distinctions, particularly in genitive constructions (A§2.2.1.4)
- the growing tendency for first-place verbs to be auxiliaries rather than main verbs
- the use in younger speakers of *tha* and *chan eil* as generalised responses for polar questions, corresponding to 'yes' and 'no' respectively (see the discussion of polar questions in 'Mode')
- use of the INDEF2 copular form only in very limited contexts, such as with 'composite verbs' (A§3.2.3)
- lack of tense agreement on the marker of comparatives and superlatives (A§2.2.4.4)
- substitution of VO order for OV (inverted nominals: see A§3.2.2) as in **feuchainn ri stad a' bhuill* 'trying to stop the ball' calqued on English; this would normally be *feuchainn ris a' bhall a stad* 'trying to the ball stop'
- blanket use of periphrastic possession with *aig* where semantic distinctions would have formerly been made, for example with inalienables (see A§2.2.1.5 and Macaulay 1982)

Code-switching is the rule in nearly all natural Gaelic conversations. No empirical study has been done on the sociolinguistic conditioning factors and syntactic consequences of code-switching in ScG (but see Cram 1981; 1986; Macaulay 1982) and it does not seem wise to offer an unsatisfying entry into this enormously complex topic at present. There is much that could be done, however, and it would be particularly interesting to view the age differences in this area. It seems reasonable that one of the first steps to researching

code-switching in ScG would be the collection of new data, particularly that garnered through naturalistic recording of informal conversations as opposed to experimental means. Although I have offered a beginning in this area, there is clearly a need for large language corpora⁹⁷ of this kind to offer a sufficiently extensive and varied data base.

4 Sample Text: Oral Folk Tale⁹⁸

bha am fear seo à Beinn na Faoghla shuas co-dhiubh agus 's e
be-PAST ART man-N this from mountain-N ART ford-G up-LOC anyway and COP 3S
'this man from Benbecula was up anyway and it is a man

fear ris an cainnte Mac a' Phì a bh' ann
man-N to COMP say-INDEF2-IMPER-DEP MacPhee REL be-PAST 'in it'
who was called MacPhee who was in it

chan cil fhios a'm dè a' chiad ainm a bh' aige co-dhiubh
be-PRES-DEP knowledge at-1S what(COP) ART first name-N REL be-PAST 'at him' anyway
I don't know what his first name was anyway

ach tha e coltach gu robh cù mòr aige agus hà cha robh
but be-PRES 3S-M seem COMP be-PAST-DEP dog-N big at-3S-M and um NEG is-PAST-DEP
but it seems that he had a big dog and um there was

sian a dh'fheum sa chù 's bhiodh a h-uile duine riamh a' ràdh ris
any of use in-ART dog-D and be-INDEF2 every man ever PROG say-VN to-3S-M
no use in the dog and every single person would say to him

ò cuir às dhan chù nach robh feum ann 's fhreagairteadh
oh put-IMP out to-ART dog-D NEG be-PAST-DEP use in-3S-M and answer-INDEF2 3S-M
oh put the dog down there no use in him and he would answer

an còmhnaidh 's e gun robh leigeadh aige fhèin aig cù Mhic a' Phì
always COP-3S-M COMP be-PAST-DEP ability at-3S-M REFL at dog-D MacPhee-G
always it's that MacPhee's dog would prove himself someday

agus co-dhiubh bliadhna a bha an-seo chaidh e fhèin
and anyway year REL be-PAST here go-PAST 3S-M REFL
and anyway this [one] year he himself

agus seachdar ghillean eile às a' bhaile chaidh iad a-mach
and seven-PN boys-G other out of ART town go-PAST 3P out-DIR
and seven other lads from out of town they went out

a chuir air dòigh nan àiridhean a-muigh a rubh' Eubhal agus an oidhche seo
to put-INF on order ART shielings out-LOC to point Eaval and ART night this
to put the shielings in order out at Eaval point and this one night

co-dhiubh nuair a bha iad deiseil dhen obair 's shuidh iad a-staigh
anyway when REL be-PAST 3P ready of-ART work-D and sit-PAST 3P in-LOC
anyway when they were done with the work and they had sat down inside

aig an teine 's thuirt na gillean ò ars' esan nach bochd a-neis
at ART fire-D and say-PAST ART boys oh quoth 3S-M-EMPH NEG-COP poor now
at the fire and the lads [sic] said oh he said isn't it too bad now

nach tigeadh seachd nighnean a-staigh 's ò ars' am bodach
NEG come-INDEF2-DEP seven girls-N in-LOC and oh quoth DET old man
that seven girls wouldn't come in and oh said the old man

⁹⁷ Around a million words would be a reasonable goal at this time.

⁹⁸ Local legend known as *Cù Dubh Mhic-a-Phì* ('MacPhee's Black Dog'), narrated by Alastair MacDonald of Croismorraig, North Uist in 1998.

is sibh a thubhairt an droch rud agus hà co-dhiubh thàinig gnogadh dhan doras
 COP 2P REL say-PAST ART bad thing and um anyway come-PAST knocking to-ART door
 isn't it you who said the bad thing and anyway there came a knocking to the door

's a-staigh a thàinig a' chailleach a bha an-seo 's seachd nìghnean
 and in-LOC REL come-PAST ART old woman-N REL be-PAST here and seven girls-N
 and in came this old woman and seven girls

's co-dhiubh 's chaidh na gillean 's na h-ighnean sìos a cheann
 and anyway and go-PAST ART boys-N and ART girls down-LOC to end
 and anyway and the lads and girls went down down to the far end

shios na h-àiridhe co-dhiubh agus hà bha am bodach
 down-LOC ART shieling-G anyway and um be-PAST ART old man-N
 of the shieling anyway and the old man was

cha robh e uabhasach toilichte 's bha e 'cumail sùil
 NEG PAST 3S-M terrible pleased and be-PAST 3S-M PROG-keep-VN eye-N
 he wasn't terribly pleased and he was keeping an eye

's seo a chunnaic e fuil a' tighinn a-nuas fon doras
 and here REL see-PAST 3S-M blood-N PROG come-VN down under-ART door-D
 and here he sees blood coming down from under the door

às a' cheann eile 's thuirt e ris a' chaillich co-dhiubh ars' esan
 out ART end other and say-PAST 3S-M to ART old woman-D anyway quoth 3S-M-EMPH
 out of the other end and he said to old woman anyway he said

tha fasan agamsa a h-uile h-oidhche a bhith a' dol a-mach a choimhead
 be-PRES fashion at-1S-EMPH every night to be-INF PROG go-VN out-DIR to look-INF
 I have the habit of going out every night to look

gu faic mi dè an t-side airson an larna-mhàireach
 to see-INDEF1-DEP 1S what-COP ART weather-N for ART morrow-G
 in order that I (might) see what the weather is for the morrow

toir sùil air an aimsir ò ars' ise mun tèid thu a-mach ma-tha
 give-VN eye-N on ART weather oh quoth 1S-F-EMPH before go-INDEF1-DEP 2S out-LOC then
 have a look on the weather oh she said before you go out then

thoir leat ceann eile dhen t-siomain tha an-siud
 take-IMP with-2S end other of-ART twisted-rope REL-be-PRES there
 take with you the other end of that rope there

'ar a bhiodh iad a-staigh air an oidhche bhiodh iad
 when REL be-INDEF-2 3P in-LOC on DAT night-D be-INDEF-2 3P
 when they used to be inside at night they would be

a' dèanamh siomain fraoich aig an teine gus a bhith aca
 PROG do-VN twisted-rope heather-G at ART fire in order to be-INF at-3P
 making a twisted heather rope at the fire so that would have it

airson an tughadh a chumail air an àirigh is gnothaichean bheir leat
 for ART thatch-N AGR keep-VN on ART shieling-D and things take-IMP with-2S
 for to keep the thatching on the shieling and things take with you

ceann an t-siomain 's cùm teann e gum bi fios agamsa
 end ART twisted-rope and keep-IMP tight 3S-M so be-INDEF1-DEP knowledge at-1S-EMPH
 the end of the rope and keep it tight so that I'll know

gum bi thu aig ceann an t-siomain 's nach teich thu air falbh
 COMP be-INDEF 2S at end ART twisted-rope and NEG flee-INDEF-DEP 2S away
 that you are at the other end of the rope and that you won't run away

agus sin a rinn e thug e a-mach ceann an t-siomain leis
 and that-COP REL do-PAST 3S-M take-PAST 3S-M out-DIR end ART rope-G with-3S-M
 and that's what he did he took out the end of the rope with him

agus cho luath 's a fhuair e a-mach air an doras cheangail e an siomain
 and as fast and REL get-PAST 3S-M out-DIR on ART door-D tie-PAST 3S-M ART rope-N
 and as fast as he got out of the door he tied the rope

leis an tughadh a bh' air bàrr na h-àiridhe 's dh'fheuch e air falbh
 with ART thatch-D REL be-PAST on top ART shieling-G and try 3S-M away
 onto the thatch that was on the top of the shieling and he ran off

dhan a' bhaile cho luath 's a bh' aige e fhèin 's an cù
 to ART town-D as fast and REL be-PAST at-3S-M 3S-M REFL and ART dog-N
 to the town as fast as he could manage - himself and the dog

agus ach co-dhiubh an ceann ùine thug e sùil as a dhèidh 's
 and but anyway in end time-G give-PAST 3S-M eye-N 'after him' and
 and but anyway in a short time he had a look behind him and

ò bha an t-each mòr geal a bha an-seo a' tighinn uca
 oh be-PAST ART horse-N big white REL be-PAST here PROG come-VN at-3P
 oh here was this big white horse coming at them

tarsainn na mònaidh ò bha an t-each mòr geal a bha seo a' tighinn uca
 across ART moor-G oh be-PAST ART horse-N big white REL was here PROG come-VN at-3P
 across the moors oh here was this big white horse coming at them

tarsainn na mònaidh aig astar eagallach as a dhèidh agus thionndaidh e ris a' chù
 across ART moor-G at speed-D frightfull 'after him' and turn-PAST 3S-M to ART dog-D
 across the moors at a frightful speed after him and he turned to the dog

's thuirt e ò ar' esan a chù dubh Mhic a' Phì mur an do rinn thu
 and quoth 3S-M oh quoth 3S-M VOC dog-V black MacPhee-G if-NEG do-PAST-DEP 2S
 and he said oh said he oh black dog MacPhee if you have never done it

a-riamh e nì thu a-nochd e agus siud an cù am bad an eich
 ever 3S-M do-INDEF 2S tonight 3S-M and that-COP ART dog-N in place ART horse-G
 before you'll do it tonight and there's the dog right at the horse

's thòisich iad ri blàr an-shin
 and begin-PAST 3P to battle-VN there-LOC
 and they began to battle there and

's fhuair esan às gu baile co-dhiubh 's dh'fhàg e
 and get-PAST 3S-M-EMPH out-3S-M to town-D anyway and leave-PAST 3S-M
 and he got away to town anyway and he left

an t-each 's an cù a' sabaid air a' mhònadh 's 'ar a thàinig e
 ART horse-N and ART dog-N PROG fight-VN on ART moor and when REL come-PAST 3S-M
 the horse and the dog fighting on the moors and when he came

dhachaigh chuir e seachd miasan bainne a-mach ri taobh an dorais
 home-DIR put-PAST 3S-M seven dishes milk-G out-DIR to side ART door-G
 home he put seven dishes of milk out beside the door

's dhùin e an doras 's nuair a dh'èirich e
and close-PAST 3S-M ART door-N and when REL rise-PAST 3S-M
 and he closed the door and when he arose

air madainn larna-mhàireach bha an cù an-shin
on morning-D next-day-G be-PAST ART dog-N there-LOC
 the next morning the dog was there

's bha e air am bainne òl 's hà cha robh sgath gaoisid
and be-PAST 3S-M after ART milk-N drink-VN and um NEG be-PAST-DEP any fur-G
 and he had drank the milk and there wasn't a bit of hair

air a' chù bha an t-each air a losgadh dheth
on ART dog-D be-PAST ART horse-N PERF 3S-POSS burn-VN off-3S-M
 on the dog the horse had burned it off him

sin agad stòiridh a' choin
there-COP at-2S story ART dog-G
 there you have the story of the dog'

Suppletive ('Irregular') Verbs

The forms given below are not meant to be exhaustive as there is significant dialectal variation with these verbs.

A. Active Voice

Root/Imperative	Verbal Noun	Form	Past	Indef. 1 'Future'	Indef. 1 Relative	Indef.2 'Conditional'
Abair <i>say</i>	ràdh(a) ràdh(a) ràitinn	<i>Ind.</i>	thuir thubhairt	abraidh their	dh'abras their	dh'abradh theirceadh
		<i>Dep.</i>	tuirt tubhairt	abair		abradh
Beir <i>catch, bear</i>	breith	<i>Ind.</i>	rug	beiridh	bheircas	bheircadh
		<i>Dep.</i>	do rug	beir		beircadh
Bi ⁹⁹ <i>be</i>	a bhith (INF only)	<i>Ind.</i>	bha	bidh	bhios bhitheas	bhiodh bhitheadh
		<i>Dep.</i>	robh	bi		biodh
Cluinn <i>hear</i>	cluinntinn cluinnteil	<i>Ind.</i>	chuala	cluinnidh	chluinneas	bitheadh chluinneadh
		<i>Dep.</i>	cuala	cluinn		cluinnceadh
Dèan <i>do, make</i>	dèanamh dèanadh	<i>Ind.</i>	rinn	nì	nì	dhèanadh
		<i>Dep.</i>	do rinn	dèan		dèanadh
Faic <i>see</i>	faicinn	<i>Ind.</i>	chunnaic	chì	chì	chitheadh
		<i>Dep.</i>	faca	faic		faiceadh
Faigh <i>find, get</i>	faighinn	<i>Ind.</i>	fhuaire	gheibh gheobh	gheibh gheobh	gheibheadh gheobhadh
		<i>Dep.</i>	d'fhuaire	faigh		faigheadh
Rach <i>go</i> thalla <i>go away</i>	dol	<i>Ind.</i>	chaidh	thèid	thèid	rachadh reidheadh dheigheadh
		<i>Dep.</i>	deach(aidh) ¹⁰⁰	tèid		rachadh d'reidheadh deigheadh
Ruig <i>arrive, reach</i>	ruigsinn ruigheachd	<i>Ind.</i>	ràinig	ruigidh	ruigeas	ruigeadh
		<i>Dep.</i>	do ràinig	ruig		ruigeadh
Thig <i>come</i>	tighinn	<i>Ind.</i>	thàinig	thig	thig	thigeadh
		<i>Dep.</i>	tàinig	tig		tigeadh
Thoir <i>take, give, bring</i>	a' toirt a' toir	<i>Ind.</i>	thug	bheir	bheir	bheireadh
		<i>Dep.</i>	tug	toir		toireadh

⁹⁹ For the present tense forms of *bi*, see A§2.3.1.2.

¹⁰⁰ In Uist, *deachaidh* is usually only used in responses and tag questions.

B. Passive Voice: Impersonal forms

These should not be confused with true passives as such: see discussion in A§2.3.2.2. To construct the impersonal form of the 2nd indefinite, some dialects (e.g. parts of Lewis) use a different suffix *-ist(e)*.¹⁰¹ For example, instead of *theirte* one finds *theirist(e)*.

Root/Imperative	Form	Past	Indef. 1 'Future'	Indef. 1 Relative	Indef.2 'Conditional'
Abair <i>say</i>	Ind.	thuirteadh thubhairteadh	abrar theircar	dh'abrar theircar	dh'abairte theirte
	Dep.	tuirteadh tubhairteadh	abrar teircar		abairte
Beir <i>catch, bear</i>	Ind.	rugadh	beircar	bheircar	bheirte
	Dep.	do rugadh	beircar		beirte
Bi <i>be</i>	Ind.	bhathar bhathas	bithear	bhithear	bhite bhithist
		robhar robhas	bithear		bite bithist
Cluinn <i>hear</i>	Ind.	chualadh chualas	cluinn(t)ear	chluinn(t)ear	chluinnte
	Dep.	cualadh cualas	cluinn(t)ear		cluinnnte
Dèan <i>do, make</i>	Ind.	rinneadh	nithear nitear	nithear nitear	dhèante
	Dep.	do rinneadh	dèantar		dèante
Faic <i>see</i>	Ind.	chunnacadh chunnacas	chithear	chitear	chite
	Dep.	facadh facas	faic(t)ear		faicte
Faigh <i>find, get</i>	Ind.	fhuaradh fhuaras	gheibh(t)ear gheobhar	gheibh(t)ear gheobhar	gheibhte
	Dep.	d'fhuaradh d'fhuaras	faigh(t)ear		faighte
Rach <i>go</i>	Ind.	chaidheas rachar	thèidear	thèidear	rachte reighte dheighte
	Dep.	deachas d'rachar	tèidear		rachte d'reighte deighte
Ruig <i>arrive, reach</i>	Ind.	ràinigeadh ràinigear ràineas	ruigear	ruigear	ruigte
	Dep.	do ràinigeadh do ràinigear do ràineas	ruigear		ruigte
Thig <i>come</i>	Ind.	thàinigeadh thàinigear	thig(t)ear	thig(t)ear	thigte
	Dep.	tàinigeadh tàinigear	tig(t)ear		tigte
Thoir <i>take, give, bring</i>	Ind.	thugadh thugas	bheircar	bheircar	bheirte thoirte thugte
	Dep.	tugadh tugas	toircar tugar		toirte tugte

¹⁰¹ The dropping of the vowel after a terminal consonant is common to the Lewis dialect.

The 100 Most Common Words of Scottish Gaelic

This list was generated by searching a 60,000 word subset of the corpus. The abbreviations employed below are glossed on page 12. By referring to the Gaelic glossary that follows, readers may find examples of these words in the text.

- 1) **a** PREP. 'to', 'at', 'into'
VOC. used to address a person or thing:
e.g. *a Sheumais* 'oh James'
3S-POSS. 'his', 'hers', 'its': e.g. *a cà* her dog; *a chù* his dog
REL. 'which', 'that', 'who', etc.
- 2) **an** ART. singular form of 'the'
3P-POSS. 'their'
- 3) **a'** ART. singular form of 'the'
- 4) **air** PREP. 'on', 'on him/it'
- 5) **tha** V. independent form of the present tense verb 'to be'
- 6) **e** 3S-M. 'him', 'he', 'it'
- 7) **na** ART. plural form of 'the'; singular form of feminine genitive: e.g. *na h-obrach* 'of the work'
REL. headless relative meaning 'all that'
- 8) **agus** CONJ. 'and'
- 9) **bha** V. past tense of verb 'to be'
- 10) **ann** PREP. 'in', 'in him/it'
- 11) **'s** CONJ. abbreviation of *agus* 'and'
V. abbreviation of the copular form of verb 'to be' *is* (present tense and 1st indefinite)
- 12) **gu** PREP. 'to'
PART. occurs before adverbs: e.g. *gu cunbhalach* 'steadily'
COMP. 'that'
- 13) **iad** 3P-M. 'them', 'they'
- 14) **ach** CONJ. 'but', 'only'
- 15) **am** QP. question particle occurring before labial consonants
ART. 'the'; occurs before singular masculine nouns beginning with labial consonants
- 16) **robh** V. dependent form of the past tense verb 'to be'
- 17) **aig** PREP. 'at'
- 18) **sin** DEM. 'that', 'there'
- 19) **gun** PREP. 'without'
COMP. 'that'
- 20) **do** 2S-POSS. 'your'
PREP. 'for', 'to'
- 21) **is** CONJ. abbreviation of *agus* 'and'
V. abbreviation of the copular form of verb 'to be' *is* (present tense and 1st indefinite)
- 22) **cha(n)** NEG. 'not'; used with independent verb forms
- 23) **nach** NEG. 'not'; used with dependent verb forms
- 24) **mi** 1S. 'me', 'I'
- 25) **i** 3S-F. 'her', 'she'
- 26) **eil** V. dependent form of the present tense verb 'to be'
- 27) **mar** PREP. 'like', 'as'
- 28) **bheil** V. dependent form of the present tense verb 'to be' (see *eil*)
- 29) **ri** PREP. 'to', 'with'
- 30) **seo** DEM. 'this', 'here'
- 31) **anns** PREP. 'in' + definite article
- 32) **de** PREP. 'of'
- 33) **chaidh** V. 'went'
- 34) **nan** ART. genitive plural definite article before non-labials: e.g. *fèis nan òran* 'festival of the songs'
CONJ. 'if'
- 35) **ag** PROG. progressive aspect marker, 'at'
- 36) **bhith** V. infinitive of the verb 'to be'
- 37) **eile** ADJ. 'other'
- 38) **le** PREP. 'with', 'by'
- 39) **fhèin** REFL. reflexive marker: e.g. *thu fhèin* 'you yourself'
- 40) **mu** PREP. 'about'
- 41) **ris** PREP. the preposition *ri* (see above) + the definite article
PREP-PN. 'to him', 'with him'
- 42) **math** ADJ. 'good'
- 43) **bhiodh** V. 'would be' (2nd indefinite form of the verb 'to be')
- 44) **b'** V. abbreviation of the copular form of verb 'to be' *bu* (past tense and 2nd indefinite)

- 45) as V. relative form of the copula (also the marker of the superlative); *an duine as fheàrr* 'the best man'
- 46) airson PREP. complex preposition 'for', 'for the sake of'
- 47) esan 3S-M-EMPH. emphatic form of the third person sing. masculine pronoun 'he', 'him'
- 48) dol VN. 'going'
- 49) aca PREP-PN. 3P-at 'at them'
- 50) thu 2S. 'you'
- 51) gum COMP. 'that' (occurs before labials)
- 52) ràdh VN. 'saying'
- 53) bho PREP. 'from', 'since'
- 54) cho PREP. 'so', 'as' (used in comparative constructions)
- 55) sinn 2P. 'we', 'us'
- 56) nuair ADV. 'when'
- 57) chuir V. past tense of *cuir* 'put'
- 58) duine N. 'man', m.
- 59) thuirt V. 'said' (past tense of *abair* 'say')
- 60) no CONJ. 'or'
- 61) f(h)ios N. 'knowledge', m.
- 62) aon NUM. 'one'
- 63) dhan PREP. the preposition *do* (see above) + the definite article
- 64) dè 'what'
- 65) leis PREP. the preposition *le* (see above) + the definite article
- 66) a-mach ADV. 'out' (directional)
- 67) dha PREP. the preposition *do* (see above) + the definite article
PREP-PN. 'to him', 'for him'
- 68) cuideachd ADV. 'too', 'also'
- 69) fear N. 'man', m.
- 70) thoirt INF. 'to give', 'to get' (from the verb *thoir*)
- 71) aige PREP-PN. 'at him'
- 72) dà NUM. 'two'
- 73) bu V. 'was'; the copular form of verb 'to be' in the past tense and 2nd indefinite
- 74) ars' V. defective verb meaning 'said', often heard in narrative
- 75) ga possessive prepositional pronoun meaning 'at his'
- 76) chur INF. 'to put' (from the verb *cuir*)
- 77) gur possessive prepositional pronoun meaning 'at your' (formal/plural)
- 78) àite N. 'place', m.
- 79) taobh N. 'side', m.
- 80) idir ADV. 'at all'
- 81) eadar PREP. 'between'
- 82) fhuair V. 'found', 'received' (past tense of *faigh* get)
- 83) co-dhiubh ADV. 'anyway'
- 84) ma CONJ. 'if'
- 85) bliadhna N. 'year', f.
- 86) nam CONJ. 'if'
ART. genitive plural definite article before labials: e.g. *fèis nam bàrd* 'festival of the poets'
- 87) uair N. 'time', f.
- 88) mòr ADJ. 'big'
- 89) tighinn VN. 'coming'
- 90) obair N. 'work', f.
- 91) agad PREP-PN. 'at him/it'
- 92) rinn V. 'did', 'made' (past tense of *dèan*)
- 93) dìreach ADV. 'exactly'
ADJ. 'straight'
- 94) thàinig V. 'came' (past tense of *tig*)
- 95) rud N. 'thing', m.
- 96) far REL. 'where' (headless relative)
PREP. 'off'
- 97) tu 2S. 'you'
- 98) bidh V. 'will be', 'is' (1st indefinite form of *bi* 'to be')
- 99) ùr ADJ. 'new'
- 100) suas ADV. 'up'

Index of Gaelic Words Used in Grammar

- 'ar*, 277
's, 254, 277
a, 200, 204, 213
à, 225
a bharrachd, 265
a bhith, 244, 251, 253, 257, 267, 274
a chùile, 216
a chionn, 258, 263
a ^{prep.}, 225
a rèir, 227
a thaobh, 227, 270
a' chiad, 219
a-bhòn-dè, 223
a-bhòn-raoir, 223
a-bhos, 224
ach, 255, 275
acras, 263
ag, 237, 244
agus, 220, 222, 253, 275, 276, 277
àidh, 272
aig, 207, 225, 243, 275
aimsir, 276
aineolach, 223
ainm, 264, 275
air, 200, 202, 225, 242
air bheulaibh, 227
air bith, 217
air chùl, 227
air dòigh, 275
air fad, 255
air falbh, 261, 276, 277
air feadh, 227
air muin, 227
air neo, 255
air sàillibh, 255, 258
air sgàth, 227, 258
air tàillibh, 258
air ^{apoc.}, 237
airgead, 249, 268, 269
àiridh, 275
airson, 217, 227, 248, 250, 256, 258, 276
àite, 217, 273, 274
aithne, 251, 255
aithris, 254
àm, 264
am bliadhna, 223
am fianais, 227
am measg, 227
a-mach, 275, 276, 277
amaideag, 203
a-màireach, 223, 235, 236, 268, 269
amais, 239
a-muigh, 247, 261, 275
an àite, 227
an ath-bhliadhna, 223
an ath-oidhche, 223
an ceart-uair, 223
an cois, 227
an còmhnaidh, 223, 275
an dèidh, 227
an earar, 223
an lùib, 266
an ^{prep.}, 225
an uiridh, 223
an/lam, 213
a-nall, 224
an-còmhnaidh, 236
an-dà, 271
an-dè, 223, 265, 268
an-diugh, 223, 249
an-dràsda, 210, 223
a-neisid, 235, 238, 275
an-fhoisil, 241
a-nìos, 224
a-nis, 223, 271
anmoch, 273
ann, 202, 221, 225, 235, 247, 265
annad, 202
an-nochd, 223
a-nochd, 231, 254, 277
an-raoir, 223
an-seo, 215, 265, 275, 277
an-shin, 277
an-sin, 215, 245
an-siud, 215, 276
a-nuas, 224, 262, 276
a-null, 224
aon, 218, 265, 270
aon uair, 262
aonan/laonar, 219
aotrom, 241
a ^{prep.}, 260
ar, 213
àrd, 203, 207, 221, 246, 255
àrd-sgoil, 203
a ^{riannaire}, 230
a-riamh, 275, 277
a-rithist, 257
ars', 276, 277
arsa, 266, 275
às, 225, 275
as aonais, 227, 241
as dèidh, 277
as ^{super.}, 221
a-staigh, 275, 276
astar, 277
a-steach, 243
at, 203
ath, 220
athair, 251
ath-athris, 203
athris, 203
atmhor, 203
bad, 277
baile, 275, 277
bainne, 206, 277, 278
ball, 243, 248, 274
bàn, 251
bàrr, 277
barrachd, 227
bàs, 273
bàsaich, 242
bàta, 220, 260
bàthaich, 208
beachd, 220
beag, 210, 220, 224, 251, 255, 262
beagan, 265
beun, 204, 210
beannachadh, 272
beò, 203
beòthail, 203
beul, 42, 240
bheir, 268
bhith, 237
bho, 249
bhon, 258
bi, 233, 234, 254, 255, 274
biadh, 255, 267
biathaich, 242
big, 220
bith-beò, 268
blàr, 277
bliadhna, 275
bò, 222, 242
bochd, 236, 261, 263, 267, 275
bodach, 205, 245, 275, 276
bodachan, 205
bòidheach, 221
boircannach, 206, 207, 265, 269
bradach, 222
bràthair, 216, 269
brist, 241
brògan, 269
brùid, 207
bruidhinn, 216, 239, 240
bu, 245, 251
buadh, 272
buaidh, 273
buannachadh, 239
buille, 258
bunrèta, 199, 273
bùrn, 42
bùth, 204, 205, 207, 229
bùth-obrach, 205
cabhag, 251
cac, 273
cadal, 262
càil, 206
caileag, 211, 212
cailleach, 276
cait, 198
càite, 216, 257, 262
call, 223, 264
can, 256, 271, 273, 275
caol, 268
caora, 219, 266
càr, 213, 235, 247, 257, 267
caraid, 220
carbad, 40
carson, 217
cas, 200, 210
cat, 40, 198, 248
ceangail, 277
ceangailte, 227
ceann, 204, 213, 258, 276
ceannach, 251
ceannard, 210
ceannsachadh, 214
ceàrdach, 210
cearr, 200
ceart-uair, 256
ceathramh, 219
ceathrar, 219
cèilidh, 43
ceithir, 218, 265
cèò, 264
cèòl, 207, 236
ceud, 218
cha(n), 230, 256
chì, 202
cho, 222, 255, 267, 277

- choireigin, 217
 chun, 225
 chunnaic, 244
 cia, 217
 ciad, 275
 ciall, 241
 ciamar, 217
 cinnteach, 257, 272
 clach, 210, 237
 cladach, 247
 cliobhar, 259
 cluin, 202, 216, 236, 253
 cnoc, 215, 228
 cnocan, 228
 cò, 217
 co-dhiubh, 266, 271, 275, 276, 277
 co-dhiin, 199
 coibhneil, 223
 còig, 218
 còigeamh, 219
 còignear, 219
 coille, 208
 coimhead, 274, 276
 còir, 238
 cois, 201
 coiseachd, 236
 coisich, 235
 coltach, 227, 275
 comasach, 239
 còmhla, 227
 compuistair, 208, 243
 còrd, 257
 corra, 220
 còta, 213, 251
 craobh, 200, 207
 creach, 272
 creachadh, 244
 creid, 241, 271
 creidsinn, 256, 272
 crios-sàbhalaigh, 273
 crodh, 206
 croir, 258
 croitear, 245
 cruthachadh, 274
 cù, 217, 218, 275, 277
 cuid, 203, 206, 268
 cuideachadh, 272
 cuideigin, 205
 cuilean, 228
 cuimhne, 272
 cuine, 217
 cuir, 232, 233, 253, 257, 261, 274, 275, 277
 cùm, 203, 259, 276
 cumail, 203, 276
 cumhang, 258
 cùramach, 203
 dà, 218
 dad, 206
 dàrna, 219
 dathach, 203
 de, 225
 de, 217, 275, 276
 deagh, 205, 220
 deamhan, 273
 dèan, 238, 240, 256, 260
 dèanadh, 203
 dèanamh, 233, 238, 242, 274, 276
 dearbh, 203, 220
 dearbhte, 203
 deas, 224
 deasaichte, 243
 deich, 218
 deicheamh, 219
 deichnean, 219
 dèidh, 237
 deireadh, 220
 deiseil, 275
 deoch, 255
 dh'fhaoite, 239
 dhà, 206
 dhachaidh, 236, 253, 277
 dhèan, 216
 dhuine, 220
 Dia, 247, 272
 diabhal, 272
 direach, 137, 224
 dithis(t), 219
 do, 213, 251, 263
 do-dhèanadh, 203
 doirbh, 266
 dol, 231, 236, 253, 276
 do^{prop}, 249
 doras, 210, 276, 277
 do-sheachanta, 241
 dragh, 253, 257
 draibheadh, 245, 257
 droch, 205, 220, 245, 273, 276
 duan, 228
 duanag, 228
 dubh, 277
 dùil, 274
 dùin, 203, 240, 277
 duine, 203, 208, 213, 238, 248, 249, 251, 263, 275
 dùthaich, 208
 e, 214
 each, 241, 265, 277, 278
 eadar, 225
 eadh, 240
 eadhon, 253, 257
 eagullach, 277
 eaglais, 249
 ear, 224
 èiginn, 223
 eile, 237, 275, 276
 èirich, 277
 èisteachd, 253, 257
 èòlach, 248
 eu-choltach, 241
 eudail, 272
 eun, 247
 fa, 223
 fada, 255
 fàg, 277
 faic, 238, 249, 251, 263, 276
 faicinn, 229, 257
 faigh, 238, 254, 255
 faighinn, 265, 270
 faisg, 227
 falbh, 240, 244, 255, 257, 258, 263, 267
 faod, 238, 239
 far, 217, 261
 fairmad, 231
 far^{prop}, 225, 239
 fàr, 203
 fasan, 276
 fastaidhear, 228
 feadhainn, 206
 fear, 204, 206, 207, 211, 217, 221, 228, 237, 273, 275
 fearan, 228
 fearg, 231, 259
 fearr, 268
 fein, 215
 fein-eachdraidh, 216
 feitheamh, 262
 feòil, 255
 feuch, 277
 feuchainn, 274
 feum, 214, 238, 274, 275
 feumaidh, 272
 fhad 's, 262
 fhalbh, 241
 fhathast, 223
 fhèin, 261, 272, 275, 277
 fhèudar, 238
 fhin, 216
 fhuair, 258
 fhead, 218
 fheadamh, 219
 fion, 231, 244
 fìor, 203, 205, 220, 245, 256
 fìor-uige, 203
 fios, 255, 257, 258, 262, 265, 272, 275, 276
 fìnghair, 274
 fo, 225, 276
 fraoch, 276
 freagarrach, 203
 fuaime, 265
 fuar, 223
 fuil, 207, 231, 276
 gabh, 262, 270, 274
 gabhail, 274
 gach, 221
 Gaidhealtachd, 43
 gàire, 232
 galla, 273
 gaoidid, 278
 geal, 204, 220, 265, 277
 ged, 259
 gille, 213, 263, 275
 giùlain, 261
 glacie, 261
 glè, 137, 223, 224, 258
 gluais, 248
 gluasad, 239
 gnogadh, 276
 gnothach, 276
 gobha, 210
 goid, 260, 262
 goirid, 227
 gonadh, 273
 gorm, 210
 gras, 272
 greis, 223
 gu, 222, 277
 gu dè, 265
 gu dearbh, 138, 224

- gu deimhinne*, 138, 224
gu lèir, 137, 224
gu math, 266
gu ruige, 227
gu siorraidh, 223
gun teagamh, 138, 224
gun/gum, 230
gun_{prep}, 225, 236, 241, 273
gu_{prep}, 225
gus, 255, 262, 276
gus_{conj}, 237
guth, 253
i, 214, 222
iad, 214
iar, 224
iarruidh, 210, 214, 255, 269
iasg, 255
iasgach, 264
ioma-dhathach, 203
iomradh, 243
ionnsaidh, 227
ire, 138, 224
is, 221, 255, 265, 273
is dòcha, 138, 224, 236, 239, 272
is mathaid, 239
is_{cop}, 245, 251
isd, 272
iteal, 223
ith, 239, 241, 242, 254, 255, 267
iuchair, 213
iùsaigeadh, 274
làmh, 203, 206, 207, 227, 265, 274
làraidh, 262
larna-mhàireuch, 276, 277
laitha, 215, 220, 238
làthair, 223
le, 204, 211, 225, 243
leabhar, 243, 247, 266
lean, 203
leantainn, 203
leigeadh, 275
leth-cheud, 218
lighiche, 208
linir, 208, 210, 238, 263
loch, 207
losgadh, 216, 278
luath, 222, 239, 277
ma, 230
ma dh'fhaoite, 138, 224
mac, 273
madainn, 277
maille, 227
mall, 40
mallachd, 273
mar, 217, 261, 264
marag, 251
marbh, 216, 220, 242
marbhadh, 231, 243, 250
mar-thà, 223, 257
mas fhìor, 272
math, 204, 224, 228, 241, 244, 258, 270
ma-tha, 271, 276
màthair, 198, 245, 263, 272
mi, 214
mias, 277
mì-chùramach, 203
mìle, 218
mìle, 218
milis, 221
milleun, 218
min, 206
mì-nàdarrach, 241
mìos, 220
mì-rùn, 203
mì-thoilich, 203
mo, 206, 213
mol, 203
moladh, 203
mòndadh, 277
mòr, 203, 205, 211, 213, 228, 255, 265, 275, 277
mòr-chuid, 203
mòr-shluagh, 205
mothaich, 222, 267
mu choinneimh, 227
mu dheidhinn, 227
mu thicheall, 227
muir, 206
mullach, 207
mun/mus, 262
mur, 277
mura, 261
mu-thrath, 223
na, 207, 217
nach, 230, 256
nàidheachd, 220
naire, 272
nan/num, 230
naoi, 218
naoidheamh, 219
naoinear, 219
na_{clausal}, 230
nas lugha, 261
nas motha, 261
neo-bhlasta, 241
neo-fhrecagarrach, 203
nighean, 220, 275
no, 255
norrag, 262
nuair, 217, 253, 262, 275
ò, 275, 277
obair, 205, 210, 213, 238, 244, 255, 275
obh, 272
och, 271, 272
ochd, 218
ochdamh, 219
ochdnar, 219
oidheche, 275, 276
oifis, 210
oir, 258
òl, 230, 231, 244, 247, 278
on, 258
òran, 205, 271
orm, 202
os, 274
os cionn, 227
piob, 207
piobair, 203
piseag, 228, 264
pòg, 216
pògadh, 251
pòsadh, 248, 250
putadh, 243
rach, 204, 230, 239, 256, 267
ràdh, 275
rathad, 239
rè, 225
reic, 265
ri, 216, 222, 225
ri linn, 227, 258
ri taobh, 277
riaghladair, 228
riaghladh, 228
riamh, 221
rian, 256
rìgh, 243
rinn, 202
ro, 137, 223, 224
ròpa, 258
ruadh, 205, 255
rubha, 275
rud, 206, 217, 265, 269, 270, 276
rùn, 203
sabaid, 259, 277
sàbhaladh, 272
sagairt, 207
sam bith, 217, 236
sàmhach, 240, 254
saoghal, 265
saoil, 202, 203
saòilsinn, 203
sàr, 220
seach, 258
seachad, 227
seachd, 218, 275, 277
seachdamh, 219
seachdar, 275
seachdnar, 219
seadh, 240, 271
seall, 202, 261
sealladh, 205
seann, 220
searbhannt, 243
sco, 215, 237, 277
seòl, 223
seòrsa, 269
sgaoil, 223
sgath, 206, 278
sgeul, 243
sgillinn, 255
sgith, 216
sgoil, 203, 210
sgriobhadh, 238
sgriobhte, 243
shuidh, 263
sia, 218
sian, 275
sianar, 219
siathamh, 219
sibh, 214
side, 276
sileadh, 231
sin, 215, 236, 237, 238, 241, 242, 249, 251, 256, 266
sinn, 214
siomain, 276
sion, 206
Siorraidh, 272
siorraidheachd, 203

- sios*, 224, 247, 276
sireadh, 210
siud, 215, 265, 269, 277
slat, 208
sluagh, 205
smaointinn, 266
snámh, 200, 228, 239
snámhaiche, 228
sneachd, 247
snog, 220
so-dhéanadh, 203
sreothart, 232
stad, 274
staighre, 261
stamh, 244
stòiridh, 278
suain, 263
suas, 224, 227, 261, 275
suidh, 275
suidhe, 231, 238
sùil, 254, 276, 277
suipear, 254
sunnd, 203
sunndach, 203
tàbh, 198
taca, 227
tachair, 253
tachairt, 274
taigh, 213, 237, 265, 273
tàmailteachadh, 199
tàmh, 198
taobh, 200, 227
tarsainn, 225, 277
tè, 206, 217, 221, 255
teann, 276
teich, 276
teine, 275, 276
telebhisean, 213
thairis, 227, 274
thall, 224
thalla, 241, 272, 273
thar, 225
theirig, 241
Thi, 272
thig, 224, 256
thogair, 272
thoir, 242, 249, 254, 267, 273
thoiri, 265
thoradh, 258
thu, 214, 269
thugainn, 241
thuir, 223
tig, 261, 275
Tighearna, 272
tighinn, 263, 265, 276, 277
timcheall, 227
tionndaidh, 277
tiugainn, 241
togail, 237
toil, 251
toilich, 203
toilichte, 257, 276
tòir, 227
toirmeasg, 251
toirt, 268, 269
tdisich, 277
treas(amh), 219
tri, 218
trioblaid, 236
triuir, 219
tro, 225
trobhad, 241
trom, 223
trusdar, 262
tuath, 224
tughadh, 276
tuigsinn, 232, 265
tuil, 207
tuilleadh, 223
tuir, 210
tunnag, 207
uabhasach, 276
uaireannan, 223
ud, 272
uell, 272
ugh, 241
Uibhisteach, 267
uile, 221, 238, 275, 276
ùine, 217, 277
uinneag, 241, 247
uisge, 42, 203
ur, 213
ùr, 264
ùrnaidh, 240
urrainn, 239, 260, 265, 271

Appendix 2: List of Texts in Corpus

Genre Name	Text Name	Words	Total	Reference (if applicable)
1. Conversation	a) NU 1	1,178	11,501	<i>All tapes recorded in North Uist between 1998 and 2000.</i>
	b) Carn	918		
	c) Truim	1,233		
	d) Truim2	528		
	e) NU 2a	1,182		
	f) NU 2a1	1,055		
	g) NU 2c	1,142		
	h) NU 2c1	1,030		
	i) NU 2c2	714		
	j) NU 4	2,521		
	<i>Genre Total</i>			
2. Public Interview	a) 'Peataichean'	2,678	10,034	Coinneach MacIomhair talk show, 1999.
	b) Fred MacAulay and Martin MacDonald	1,049		Biographic television show on Fred Macaulay, rebroadcast by Radio nan Gaidheal in 1998.
	c) McInnes and Matheson	1,094		Rebroadcast of discussion between Celtic scholars John MacInnes and William Matheson (originally from 1970s?).
	d) Gèamaichean Shollais	1,206		Live coverage of 2000 Highland Games in North Uist, as part of Coinneach MacIomhair's show.
	e) G Shollais II	1,154		Live coverage of 2000 Highland Games in North Uist, as part of Coinneach MacIomhair's show.
	f) Bonn Comhraidh	1,145		Political discussion show from the 1980s.
	g) CM 17 Jan 2000	1,025		Coinneach MacIomhair talk show, 2000.
	h) CM 17 Jan II	953		Coinneach MacIomhair talk show, 2000
	<i>Genre Total</i>			
3. Sports Reportage	a) Alba vs. Astràilia 1	1,001	10,278	<i>Notes: tapes of Radio nan Gaidheal coverage of two football games broadcast in the early 1990s.</i>
	b) Alba vs. Astràilia 2	986		
	c) Alba vs. Astràilia 3	991		
	d) Alba vs. Astràilia 4	959		
	e) Alba vs. Astràilia 5	1,035		
	f) Alba vs. Yugoslavia 1	1,132		
	g) Alba vs. Yugo 2	1,045		
	h) Alba vs. Yugo 3	1,098		
	i) Alba vs. Yugo 4	1,007		
	j) Alba vs. Yugo 5	1,024		
	<i>Genre Total</i>			
4. Oral Narrative	a) Na Tri Leinntean Canaich	1,066		Johnson, D.A. 'Na Tri Leinntean Canaich', SA 1971/42 and 43.
	b) Conall Gulban	1,002		

	c) Na Fiantaichean	996	McLellan, Angus 'Mar a bha an t-sealg air a cleith air na Fiantaichean', SA 1963/15/B.
	d) Gille an Fheadain Duibh	1,051	Moiriasdan, Pàdruig (1977) 'Gille an Fheadain Duibh', in Domhnallach, D.E. (ed.) <i>Ugam agus Bhuam</i> , 1-3, Steornabhagh: Club Leabhar, SA 1967/118/B1 and SA 1966/82/A2-fB1.
	e) Bodach Ròcabarraigh	1,123	Moiriasdan, Pàdruig (1977) Bodach Ròcabarraigh', in Domhnallach, D.E. (ed.) <i>Ugam agus Bhuam</i> , 53-56, Steornabhagh: Club Leabhar, SA 1972/20/A2.
	f) Iain Beag MacAnndra	828	MacLean, Mary Margaret 'Iain Beag MacAnndra', SA 1971/3/A5.
	g) Fear a' Churraicain Ghlais	1,008	MacDougal, Donald 'Fear a' Churraicain Ghlais' SA1968/212/B5 and 213/A1.
	h) Boban Saor	1,036	MacDougal, Donald 'Boban Saor agus a Mhac', SA 1971/183/B5 and 184/A1.
	i) Bean 'ic Odrum	1,068	MacDougal, Donald 'Bean Mhic Odrum', SA 1968/212/B1.
	j) Blàr Chàirinis	1,043	Robertson, Calum 'Blàr Chàirinis', SA 1968/273/A4.
	<i>Genre Total</i>		10,221
5. News Scripts	a) Màiri Anna NicUalraig 1	684	<i>All scripts from late 1990s. Sent to author over e-mail from workers at Radio nan Gaidheal.</i>
	b) Dòmhnall Moiriasdan 1	569	
	c) Iseabail NicIllinnein	1,094	
	d) Innes Rothach	1,878	
	e) Innes Rothach 2	439	
	f) Pàdraig MacAmhlaigh	624	
	g) Dòmhnall Moiriasdan 2	1,019	
	h) Màiri Anna NicUalraig 2	1,916	
	i) Seumas Domhnallach	1,005	
	j) Seumas Domhnallach 2	1,006	
	<i>Genre Total</i>		10,234
6. Fiction	a) Am Fainne	2,314	Watt, Eilidh (1998) 'Am Fainne' in <i>Starsach: Rosg is Bàrdachd</i> , 51-56, Highland Printers: Inbhir Nis.
	b) Cùmhnantan, Chapt. 1	1,037	MacGill-Eain, Tormod (1996) <i>Cùmhnantan</i> , Clò Loch Abair: Glaschu.
	c) Droch Àm	1,029	MacAonghais, Pòl (1979) 'Droch am dhe'n bhliadhna' in <i>Amannan: Sgialachdan Goirid</i> , 1-16, Dùn Èideann: MacGillemhòire and Gibb Teo.
	d) Spàl Tim	1,164	MacCoinneach, Cailean T. (1998) 'Spàl Tim' in <i>Starsach: Rosg is Bàrdachd</i> , 4-12, Highland Printers: Inbhir Nis.

	e) Teine a Loisgeas	999	Watt, Eilidh (1979) 'Teine a Loisgeas' in <i>Amannan: Sgialachdan Goirid</i> , 17–27, Dùn Èideann: MacGillemhóire and Gibb Teo.
	f) Beul na h-Oidhche	1,022	Mac Gill-eain, Somhairle (1997) 'Beul na h-Oidhche' in William Gillies (ed.), <i>Ris a' Bhruthaich: The Criticism and Prose Writings of Sorley MacLean</i> , 267–272, Acair: Stornoway.
	g) An t-Aonaran chapt 1	1,055	Mac a' Ghobhainn, Iain (1976) <i>An t-Aonaran</i> , Gairm: Glaschu.
	h) Briseadh na Cloiche	1,010	Moireach, Iain (1985) 'Briseadh na Cloiche' in <i>Eadar Pèann is Pàipear</i> , Gairm: Glaschu.
	<i>Genre Total</i>		9,630
7. Formal prose	a) Tri Ginealaichean	1,837	Dòmhnallach, D.E. (1989) 'Tri ginealaichean de sgeulachd', in W. Gillies (ed.), <i>Gaelic and Scotland</i> , 185–221.
	b) Nua-Bhàrdachd Ghàidhlig	1,009	MacAmhlaigh, Dòmhnall (1976) 'Roimh-radha' in <i>Nua-Bhàrdachd Ghàidhlig</i> , Canongate: Edinburgh.
	c) Mairead N. Lachlainn	1,002	Mac Gill-eain, Somhairle (1997) 'Beul na h-Oidhche' in William Gillies (ed.), <i>Ris a' Bhruthaich: The Criticism and Prose Writings of Sorley MacLean</i> , 162–176, Acair: Stornoway.
	d) Bith-eòlas	1,069	MacThòmais, R. (1976) <i>Bidh-Eòlas: A' Chealla, Ginntinneachd, is Mean-fhas</i> , Gairm: Glaschu.
	e) Aramach am Bearnaraidh	1,041	In MacAmhlaigh, Domhnall (ed.) (1983?) <i>Oighreachd agus Gabhaltas: Iomraidhean air Aimeireit an Fhearrainn anns na h-Eileanan</i> , Edinburgh: Macdonald Printers.
	f) Blàr a' Chumhaing	1,007	In MacAmhlaigh, Domhnall (ed.) (1983?) <i>Oighreachd agus Gabhaltas: Iomraidhean air Aimeireit an Fhearrainn anns na h-Eileanan</i> , Edinburgh: Macdonald Printers.
	g) Na Marbhrannan	973	MacDhòmhnaill, Coinneach D. (1989) 'Na Marbhrannan Soisgeulach' in W. Gillies (ed.), <i>Gaelic and Scotland</i> , 175–184.
	h) Caint is Cànan	1,002	MacInnes, J. (1995?) 'Caint is Cànan', <i>An Tarbh</i> , Sabhal Mòr Ostaig.
	i) Trachdas DW	1,149	Stiùbhart, Domhnall Uilleam (1999) unpublished PhD dissertation, University of Edinburgh.
	<i>Genre Total</i>		10,089
8. Popular Writing	a) An Cuir am Papa...	1,060	O Hianlaidh, Aileig, (1999) 'An cuir am Papa aonta ri easbaig le Gàidhlig?', <i>An t-Albannach</i> , 10/3/1999.

b) A bhith mar Chorra...	673	NicDhomhnaill, Joina. (1999) 'A bhith mar chorra-ghrithreach air sliabh na beinne', <i>An t-Albannach</i> , 26/2/1999.
c) Pàdraig Sellar	805	MacIlIinnein, Ùisdean (1999) 'Seallair air beulaibh na cuirte – a-rithist', <i>An t-Albannach</i> , 3/11/99.
d) A' Cur Às Dhuinn Fhìn	860	Mac-a-Phì, Aonghas (1999) 'A' cur às dhuinn fhìn le faoineas 's cion-doighe', <i>An t-Albannach</i> , 10/8/99.
e) Aon Dùthaich	870	MacLeòid, Murchadh (1999) "'Aon dùthaich" le cothrom dhan Ghàidhlig?', <i>An t-Albannach</i> , 22/9/99.
f) Blas a' Ghuga	697	MacLeòid, Coinneach (1999) 'Blas a' ghuga do mhuinntir na h-Eireann', <i>An Gàidheal Ùr</i> , web cite, http://an-gaidheal-ur.co.uk/four.html .
g) Luchd-ciùil	1,096	Dick, Crìosaidh (1999) 'Luchd-ciùil a' cluich air son an toil fhèin', <i>An t-Albannach</i> , 17/09/99.
h) Na Gaidheil Ùra	1,048	Dick, Crìosaidh (1999) 'Gàidhlig fhileanta 's gun iad nan Gaidhil', <i>An t-Albannach</i> , 19/11/99.
i) A' Siubhail gu Rèidh	761	Domhnallach, Tormod (1999) 'A' siubhail gu rèidh bho eilean gu eilean', <i>An t-Albannach</i> , 24/11/99.
j) Poileaticeans	1,036	Brownlie, Niall M. (1999) 'Poileaticeans cho caochlaideach ri taobh na mara' <i>An t-Albannach</i> , 29/10/99.
g) Oifigeir Gàidhlig	784	O Hianlaidh, Aileig (1999) 'Oifigeir Gàidhlig air son Holyrood', <i>An t-Albannach</i> , 10/11/99.
<i>Genre Total</i>		9,690
Total:		81,677

Appendix 3: Full Tag Set

I. Clausal Features

Information Structure, Focus, and Discourse

Tag	Feature Name	Examples
FC(A)*	Cleft Construction: Fronting Implicated	's e fhèin a chaidh ann; <i>cha mhise a rinn e</i> 's ann a chaidh e fhèin ann; <i>chan ann an-shin a bha e</i>
FD	Demonstrative Copular Structure	('s e) <i>siud</i> an duine a rinn e; <i>sin</i> thu fhèin
FI	'Identificational' Clefts	's e fear Uibhisteach a bh' ann; 's e buille lag a bha sin
FO(A)*	Other Copular Constructions	is duine còir e; 's mise an dotair an-seo; is breàgha an tè ud
FP	'Pseudo clefts' and other cleft-like structures	an rud a th' ann 's e ...; chan eil ann ach rud gun rud
FE	Centre-detached Syntagms	tha <u>iad</u> an nighean 's an gille a' tighinn dhachaigh an-diugh
FF	Fronted adverbials and peripheral elements	<i>a-mach a ghabh iad</i> ; <i>os cionn na cidsin</i> tha leapanan ann
FL	Left-detached elements	<i>Seumas dh'fhalbh e</i> mar-thà; <i>an taigh ud</i> tha e grànda
FN	Pre-core nominal elements	ach <i>Seumas</i> cha ghabhadh am biadh
FR	Right-detached elements	tha e grànda <i>an taigh ud</i>
FX	Repetition and expansion of clause	bha iad math bha iad uabhasach fhèin math
e	Verbal Elipsis	<i>hovel</i> uabhasach a th' ann

*--(A) indicates that the fronted element is of the *ann* type – generally adjectival, adverbial

Clause Types

Tag	Feature Name	Examples
CA	<i>Agus-</i> and <i>ach-</i> headed co-subordinate clause	thuit e còrr is ceud troigh 's e a' sreap air Aonach Dubh
CCA	Complement clause after adjective or non-verbal elements	tha mi <u>toilichte</u> gun tàinig thu; bha <u>agal</u> air nach robh sinn ga thuigsinn
CCV	Complement clause after verbal elements	cha <u>chreid</u> mi gum faca e mi; tha mi a' <u>smaointinn</u> gun creid sibh seo
CF	Fragment (e.g., short responses in dialogue)	seadh; bhà; cha deachaidh
CFS	False Starts – speaker changes direction, rephrases	<i>uell tha iad</i> tha an t-airgead a tha iad a' cosg gun fheum
CM	Main clause	dh'iarr mi orra...; an robh e ann?; sguir
CN	Clausal negation	chan eil...; nach bi...
CO	Coordinate clause (clause is independent)	...agus cha tuirt mi smid; ...ach cha robh e ann
CON	Clausal conjunction (clause is not independent)	thuit iad ris gun robh e ann <i>agus gum faca e am bodach</i>
CQ	Direct Interrogative Clause	an tuirt thu rudeigin?; dè bha e 'deanamh?
CT	Terminal Tag	tha e math <i>nach eil?</i>

Subordinate Clause Types (takes either F or P as affix on tag (fronted/ preposed))

Tag	Feature Name	Examples
CSC_	Cause and Reason	a chionn; air sgàth; a thoradh; bhon
CSN_	Conditional	nan robh...; mur am bi...; ma tha...
CSO_	Other	ge brith dè; cho doirbh 's nach urrainn mi
CSS_	Concessive	ged
CSV_	Temporal	nuair; fhad 's; gus; mus; bho

Relative Clauses and Noun Phrase Complexity

Tag	Feature Name	Examples
CR	Relative Clause	an duine a bha ann; na bh' againn de bhiadh
CRC	Copular RC	am fear a b' aithne dhomh
CRP	Possessive RC	an nighean a bha a màthair bochd
CRR	Prepositional RC	tha creag mhòr ris an cainnte...
NC	Complex Noun Phrase – has 2 or more modifiers	fear-gairm nan Lib Deamach 'sa Phàrlamaid; duilgheasdan teicneigeach is eile; seann taigh geal

Valence Decreasing Constructions – can take (A)gent as affix (for when an agent is present)

Tag	Feature Name	Examples
PA_	Analytic Passives and Impersonals with <i>dol</i>	thèid a dheanamh; chaidh a cur air bhog le Rùnaire na Stàite
PO_	Other Valence Decreasing Constructions	tha dùil; tha fhios
PS_	Synthetic 'Passives' and Impersonals	thogadh mi; gum faicear; rinneadh e; chainnte
PV_	Valence Decreasing Encoded with <i>air</i> , <i>ri</i> , and <i>aig</i>	bha i 'ga cumail an-sin; bidh iad gam faicinn

Miscellaneous

Tag	Feature Name	Examples and Definitions
•	Construction of Interest	rare and interesting linguistic specimens
INC	Incomplete Clauses	transcription difficulties, poetry, non-native speaker, etc. <i>Items Not to Be Scored</i>
ENG	English Code Switch	<i>but that doesn't matter Ailean</i> ma tha an taigh glan

II. Word-Level Features

Verbal Morphology: (I)ndependent or (D)ependent as affix to indicate verbal status

Tag	Feature Name	Examples
VC_	Conditional-Habitual/ Past-tense Copula	bhiomaid; bhiodh e; b' àbhaist dha; cha bu mhise
VF_	"Future"/ Non-past Indefinite	bidh iad; ruithidh; cha bhi
VI	Imperative and Subjunctive	isd thusa; bithibh sunndach; na cluinneam
VL	Infinitival Complement	chaidh mi a <i>shealltainn</i> ; a bheil thu airson a <i>chluinntinn</i>
VN	Verbal Noun	ag iarraidh; 'ga dheanamh; 'nam shuidhe
VP_	Past Tense	bha; chaidh; rinn
VPP	Past Participle*	deanta; roinnte; sgaraichte
VT_	Present Tense and Present Copula	is mise; tha; a bheil
COP	Copula Verb	is; bu; <i>as</i> aithne dhomh
CV	"Composite" Verb	bu mhath leam; bu chòir dha; eg., <i>is</i> <VTI_COP> aithne <CV> <i>dhomh</i> <RP>

*--these are rather rare as such in ScGaelic and many of them could be better characterised as a class of adjectives

Aspectual Marking

Tag	Feature Name	Examples
SF	Perfective: <i>air</i>	tha mi air sin a dheanamh
SP	Progressive: <i>ag</i>	tha mi a' deanamh cus
SS	Prospective and Predicative: <i>ri, gus, an impis, a' dol</i> a INF, <i>gu bhith</i>	tha e ri iasgach; chan eil sin ri fhaighinn

Nominal System

Tag	Feature Name	Definition
M/F/E	Gender/English Noun	Masculine; Feminine; English Noun
S/P	Count	Singular; Plural
D/I	Definiteness	Definite; Indefinite
V/N/D/G	Case	Vocative; Nominative; Dative; Genitive
C/U/A	Grammatical Concordance	Concordant; Unconcordant; Assimilated <i>This feature refers to whether or not the lexical item in question is formed in concordance with the parameters set out in my reference grammar. The assimilation feature indicates, in cases of English borrowing, whether the lexical item has been gaelicised to an extent.</i>

Syntax for Nominal System Tags: Gender – Count – Definiteness – Case – Gram. Conc.

Ex.: MSINC

Proper Nouns

Tag	Feature Name	Definition
N/E	Proper Noun Tag	Name; English*
V/N/D/G	Case	Vocative; Nominative; Dative; Genitive
C/U/A	Grammatical Concordance	Concordant; Unconcordant; Assimilated

*- Although this same tag is used with common nouns, the syntax is different here:

Ex.: NDC; ENA

Parenthesis, Possession, Pied-piping, Repetition and Expansion

Tag	Feature Name	Examples
+ ___ +	Parenthetical Discourse	fhios agad; mar a bha mi a' ràdh; tha mi a' creidsinn
PD	Demonstrative Pronoun	seo; sin; siud
PN	Pronoun	e; i; etc.
PNI	Indefinite Pronoun Marking	-eigin; air choreigin; sam bith
PNR	Resumptive (hanging) Pronoun	an duine a bha mi a' bruidhinn ris
PP	Periphrastic Possession	an càr aig...
PPN	Possessive Pronoun	a bhean
R(1,2)	Compound (R1) and Complex (R2) Prepositions	comhla ri <R1>; mu choinneimh <R2>
RE	Repetition and Expansion of Phrasal Elements	feumaidh rùm a bhith ann rùm mòr a bhith ann
RP	Prepositional Pronoun	dhaibh; orm; etc.

Adverbs and Adjectives

Tag	Feature Name	Examples and Definitions
A-C/U-(P)	Attributive Adjective	Concordance has the same definition as above P refers to pre-nominal position, as in adjectives such as <i>deagh ex.</i> , duine math <AC>; duine mhath <AU>; deagh <ACP> dhuine
AD	Predicative Adjective	tha an duine math <AD>
AP	Place Adverbial	a-nall; air ais; a-nuas; an-sin
AS	Superlatives and Comparatives	na bu treasa; as soilleire
AT	Time Adverbial	an-diugh; fhathast; a-nisd
AV	Adverb	gu math; gu luath; seachad; air adhart

Miscellaneous Categories and Lexical Classes

Tag	Feature Name	Examples and Definitions
LA	Amplifier	gu lèir; gu mòr; buileach; uabhasach; ro; gu h-iomlan; direach; gu math; gu tur; glè – <i>gives some indication of degree</i>
LB	Subordinator	ma; nam; o'n; oir; a chionn; air sàilleabh
LC	Conjunct	a bharrachd; a-rèisde; githheadh; air an làimh eile; ge-tà; mar sin dhe – <i>refer to clausal relations and occur generally at beginning of clause</i>
LD	Downtoner	gu ìre mhòir/ bhig; gu beag(naich) – <i>some indication of degree of uncertainty</i>
LE	Emphatic	gu dearbh; gu deimhinne; gun teagamh; gu cinnteach – <i>just presence of certainty, no degree of which is given</i>
LF	Focus Clitic	-sa, -se etc.
LG	Agus etc. in comparative and subordinate-type constructions	a chionn 's; cho fad agus a tha
LH	Hedge	cha mhòr; mu; 's dòcha; ma dh'fhaite; 's mathaid – <i>less-specific indications of diminishing probability or uncertainty</i>
LJ	Demonstrative Adjective	an duine sin
LK	Numerals	aon; a' chiad ¹⁰²
LL(A)	English Borrowing	For non-nominal lexical items (eg., adjectives or verbs) (A)ssimilated as above

¹⁰² Numerative pronouns such as *dithis*, and *sianar* are treated as nouns.

LN	Synthetic Negation	gun; as aonais
LO	Conjunction	no; agus; ach.
LP	Public Verb	innse; aideachadh; gearain; mineachadh; ràdh
LR	Discourse Particle	àidh; uell; co-dhiubh; a-nisd – <i>maintain conversational coherence</i>
LS	Suasive Verb	thoir gu creidsinn; a' moladh; iarraidh; a' toirt air
LV	Private Verb	smaoineachadh; saoilinn; beachdachadh
LX	Reflexives and Reciprocals	fhèin; a chèile
LZ	Appositive	thuirt Mo Mowlam rùnaire Èirinn-a-Tuath...

References

- Adger, David (1997) 'VSO order and weak pronouns in Goidelic Celtic', *Canadian Journal of Linguistics*, 42(1–2): 9–29.
- Akinnaso, F. Niyi (1985) 'On the similarities between spoken and written language', *Language and Speech*, 28: 323–59.
- Atkinson, Dwight and Biber, Douglas (1994) Register: A review of empirical research', in D. Biber and E. Finegan (eds.), 351–85.
- Baker, P., Lie, M., McEnery, T., and Sebba, M. (2000) 'The construction of a corpus of spoken Sylheti', *Literary and Linguistic Computing*, 15(4): 421–31.
- Beaman, Karen (1984) 'Coordination and subordination revisited: Syntactic complexity in spoken and written narrative discourse', in D. Tannen (ed.) *Coherence in Spoken and Written Discourse*, 45–80, New Jersey: Ablex.
- Besnier, Niko (1988) 'Spoken and written Nukulaelae registers', *Language*, 64(4): 707–36.
- Besnier, Niko (1998) Review of Biber, D. (1995), *Language in Society*, 27(1): 126–9.
- Biber, D. (1986) 'Spoken and written textual dimensions in English: resolving the contradictory findings', *Language*, 62: 384–414.
- Biber, D. (1988) *Variation Across Speech and Writing*, Cambridge and New York: Cambridge University Press.
- Biber, D. (1990) 'Methodological issues regarding corpus-based analyses of linguistic variation', *Literary and Linguistic Computing*, 5(4): 257–69.
- Biber, D. (1994) 'An analytical framework for register studies', in D. Biber and E. Finegan (eds.) *Sociolinguistics Perspectives on Register*, 31–56, New York: Oxford University Press.
- Biber, D. (1995) *Dimensions of Register Variation: A Cross-linguistic Comparison*, Cambridge: Cambridge University Press.
- Biber, D. (1998) *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge: Cambridge University Press.
- Biber, D. and Finegan, E. (1994) 'Introduction: Situating register in sociolinguistics', in D. Biber and E. Finegan (eds.), 3–12.
- Biber, D. and Finegan, E. (eds.) (1994) *Sociolinguistics Perspectives on Register*, New York: Oxford University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999) *The Longman Grammar of Spoken and Written English*, Harlow, UK: Longman.
- Blankenship, Jane (1962) 'A linguistic analysis of oral and written style', *Quarterly Journal of Speech*, 48: 419–22.
- Blankenship, Jane (1974) 'The influence of mode, submode, and speaker predilection on style', *Speech Monographs*, 41: 85–118.
- Bloomfield, Leonard (1935) *Language*, New York: Holt, Rinehart, and Winston.
- Borgström, C.H.J. (1937) 'The Dialect of Barra in the Outer Hebrides', *Norsk Tidsskrift for Sprogvidenskap*, 8: 71–242.
- Borgström, C.H.J. (1940) *The Dialects of the Outer Hebrides: A Linguistic Survey of the Gaelic Dialects of Scotland* vol I., Norwegian Universities Press.
- Borgström, C.H.J. (1941) 'The Dialects of Skye and Ross-shire', *Norsk Tidsskrift for Sprogvidenskap* supplementary vol. I.
- Boyle, D. (1973) 'Ach and agus as coordinate and subordinate conjunctions in Gaelic', in Corum, C., T.C. Smith-Stark and A. Weiser (eds.) *You take the high node and I'll take the low node*, 220–28, Chicago: Chicago Linguistics Society.

- Brown, G., and Yule, G. (1983) *Discourse Analysis*, Cambridge: Cambridge University Press.
- Bruford, Alan (1969) *Gaelic Folk-tales and Mediaeval Romances*, Dublin: Dublin Institute for Advanced Studies
- Bruford, Alan (1978) 'Recitation or re-creation: examples from South Uist storytelling', *Scottish Studies*, 22.
- Burnard, Lou (1995) 'The Text Encoding Initiative: an overview', in G. Leech, G. Myers, J. Thomas (eds.), 69–81.
- Byrne, Michel (2000) *Facal air an Fhacal*, SRG – Acair Earranta: Steòrnabhagh.
- Calder, G. (1923) *A Gaelic Grammar* (1990 edition), Glasgow: Gairm Publications.
- Campbell, John F. (1994) *Popular Tales of the West Highlands*, Edinburgh: Birlinn.
- Campbell, John L. (ed.) (1996) 'The Rev. Fr. Kenneth MacLeod's collection of Gaelic asseverations, exclamations, and imprecations', *Scottish Gaelic Studies*, 14: 71–81.
- Chafe, Wallace (1982) 'Integration and involvement in speaking, writing and oral literature', in D. Tannen (ed.) *Spoken and Written Language: Exploring Orality and Literacy*, 35–53, Norwood, NJ: Ablex.
- Chafe, Wallace (1984a) 'How people use adverbial clauses', *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*, Berkeley: University of California.
- Chafe, Wallace (1984b) 'Speaking, writing, and prescriptivism', in D. Shiffrin (ed.), 95–104.
- Chafe, Wallace (1985) 'Linguistic differences produced by differences between speaking and writing', in Olson, D.R., Torrence, N., and A. Hildyard (eds.) *Literacy, Language, and Learning*, Cambridge: Cambridge University Press.
- Chafe, Wallace (1987) 'Cognitive constraints on information flow', in R.S. Tomlin (ed.) *Coherence and Grounding in Discourse*, 21–51, Amsterdam: John Benjamin.
- Chafe, Wallace (1988) 'Punctuation and the prosody of written language', *Written Communication*, 5: 396–426.
- Chafe, Wallace (1992) 'The importance of corpus linguistics to understanding the nature of language', in Svartvik, J. (ed.), 79–97.
- Chafe, W. and Danielewicz, J. (1987) 'Properties of written and spoken language', in R. Horowitz and S.J. Samuels (eds.) *Comprehending Oral and Written Language*, 83–113, New York: Academic Press.
- Chafe, W. and Tannen, D. (1987) 'The relation between written and spoken language', *Annual Review of Anthropology*, 16: 383–407.
- Collins, Peter C. (1991) *Cleft and Pseudo-cleft Constructions in English*, London: Routledge.
- Collins, Peter and Lee, David (eds.) (1999) *The Clause in English*, Amsterdam: John Benjamins.
- Constantinidou, Evi (1994) 'The "Death" of East Sutherland Gaelic: Death by Women?', in P. Burton, K. Dyson and S. Ardener (eds.) *Bilingual Women: Anthropological Approaches to Second-Language Use*, 111–27, Oxford and Providence: Berg.
- Coulthard, R. and Robinson W. (1968) 'The structure of the nominal group and elaboratedness of code', *Language and Speech*, 11(4): 234–50.
- Cox, Richard A.V. (1998) 'Tokenism in Gaelic: The language of appeasement', *Scottish Language*, 17: 70–81.
- Cram, David (1981) 'Code-switching, pidgin Gaelic and *Whisky Galore*', *Scottish Gaelic Studies*, 13: 241–62.

- Cram, David (1984) 'Is Scottish Gaelic a VSO language?', *Scottish Language*, 3: 33–48.
- Cram, David (1986) 'Patterns of English-Gaelic and Gaelic-English switching', *Scottish Language*, 5: 126–30.
- Crowdy, S. (1993) 'Spoken corpus design', *Literary and Linguistic Computing*, 8(4): 259–65.
- Crowdy, S. (1995) 'The BNC spoken corpus', in G. Leech, G. Myers, and J. Thomas (eds.), 224–34.
- Dieckhoff, H.C. (1932) *A Pronouncing Dictionary of Scottish Gaelic*, Edinburgh and London.
- Dòmhnallach, D.E. (1989) 'Trì ginealaichean de sgeulachd', in W. Gillies (ed.), 185–221.
- Dorian, Nancy (1981a) *Language Death: The Life Cycle of a Scottish Gaelic Dialect*, Philadelphia: University of Pennsylvania Press.
- Dorian, Nancy (1981b) 'The valuation of Gaelic by different mother tongue groups resident in the Highlands', *Scottish Gaelic Studies*, 13: 169–82.
- Dorian, Nancy (1986) 'Making do with less: Some surprises along the language death proficiency continuum', *Applied Psycholinguistics*, 7: 257–76.
- Dorian, Nancy (1994a) 'Stylistic variation in a language restricted to private-sphere use', in D. Biber and E. Finegan (eds.) *Sociolinguistics Perspectives on Register*, 217–34, New York: Oxford University Press.
- Dorian, Nancy (1994b) 'Varieties of variation in a very small place: Social Homogeneity, prestige norms, and linguistic variation', *Language*, 70(4): 631–96.
- Dorian, Nancy (1999) 'Celebrations: In praise of the particular voices of languages at risk', *Ogmios*, 12: 4–14.
- Dressler, Wolfgang U. (1988) 'Language death', in F. Newmeyer (ed.) *Linguistics: The Cambridge Survey IV*, 184–92, Cambridge: Cambridge University Press.
- Durkacz, V.E. (1983) *The Decline of the Celtic Languages*, Edinburgh: John Donald Publishing.
- Dwelly, Edward (1901–1911) *The Illustrated Gaelic-English Dictionary* (1993 edition), Birlinn: Edinburgh.
- Ferguson, C. (1968) 'Language development', in J. Fishman, C. Ferguson, and J. Gupta (eds.) *Language Problems of Developing Nations*, 27–35, New York: John Wiley and Sons.
- Ferguson, C. (1983) 'Sports announcer talk: Syntactic aspects of register variation', *Language in Society*, 12: 153–72.
- Ferguson, C. (1994) 'Dialect, register, and genre: Working assumptions about conventionalization', in D. Biber and E. Finegan (eds.), 15–30.
- Foley, W. and Van Valin, R. (1986) *Functional Syntax and Universal Grammar*, New York: Cambridge University Press.
- Fox, Barbara and Thomson, Sandra (1990) 'A discourse explanation of the grammar of relative clauses in English conversation', *Language* 66(2), 297–316.
- Garside, Roger (1995) 'Grammatical tagging of the spoken part of the British National Corpus: a progress report', in G. Leech, G. Myers, J. Thomas (eds.), 161–7.
- Ghadesy, Mohsen (1988) 'The language of written sports commentary – Soccer: A description', in M. Ghadesy (ed.) *Registers of Written English: Situational Factors and Linguistic Features*, 17–51, London: Pinter Publishers.
- Gillies, W. (1980) 'English influences on contemporary Scottish Gaelic', *Scottish Literary Journal, Supplement 12 (Language)* 1–12.

- Gillies, W. (1987) 'Scottish Gaelic: the present situation', in G. MacEòin, A. Ahlqvist and D. Ó hAodha (eds.) *3rd International Conference on Minority Languages: Celtic Papers*, 27–46, Clevedon: Multilingual Matters.
- Gillies, W. (ed.) (1989) *Gaelic and Scotland*, Edinburgh: Edinburgh University Press.
- Gillies, W. (1993) 'Scottish Gaelic', in Martin Ball (ed.) *The Celtic Languages*, 145–227, London: Routledge.
- Gillies, W. and Matheson, W. (1994) 'Gaelic personal names', in D. Thomson (ed.) (1994a), 210–1.
- Givón, Talmy (ed) (1979) *Syntax and Semantics* Volume 12: *Discourse and Syntax*, London: Academic Press.
- Gleasure, J.W. (1994) 'Gaelic: dialects, principal divisions', in D. Thomson (ed.) (1994a), 91–5.
- Goutsos, D., Hatzidaki, O., and King, P. (1994) 'Towards a corpus of spoken modern Greek', *Linguistic and Literary Computing*, 9(3): 215–23.
- Green, David (1994) 'Gaelic: syntax, similarities with British syntax', in D. Thomson (ed.) (1994a), 107–8.
- Greenbaum, S. and Nelson, G. (1995a) 'Clause relationships in spoken and written English', *Functions of Language*, 2: 1–21.
- Greenbaum, S. and Nelson, G. (1995b) 'Nuclear and peripheral clauses in speech and writing', in G. Melchers and B Warren (eds.) *Studies in Anglistics*, 181–90, Stockholm: Almqvist and Wiksell.
- Greenbaum, S. and Nelson, G. (1996) 'Positions of adverbial clauses in British English', *World Englishes*, 15: 699–781.
- Greenbaum, S. and Nelson, G. (1999) 'Elliptical clauses in spoken and written English', in P. Collins and D. Lee (eds.), 111–26.
- Greenbaum, S., Nelson, G. and Weitzman, M. (1996) 'Complement clauses in English', in J. Thomas and M. Short (eds.) *Using Corpora for Language Research: Studies in the Honour of Geoffrey Leech*, 76–91. London: Longman.
- Haiman, J. and Thompson, S. (1985) "'Subordination" in Universal Grammar', in Brugman *et al.*, *Proceedings of the Berkeley Linguistics Society*, 510–21.
- Halliday, MAK (1987) 'Spoken and written modes of meaning', in Rosalind Horowitz and S. Jay Samuels (eds.) *Comprehending Oral and Written Language*, 55–82, New York: Academic Press.
- Halliday, MAK (1989) *Spoken and Written Language*, Oxford: Oxford University Press.
- Halliday, MAK (1992) 'Language as system and language as instance: The corpus as a theoretical construct', in Svartvik, J. (ed), 61–78.
- Halliday, MAK (1994) *An Introduction to Functional Grammar*, London: Arnold.
- Hamp, E. (1970) 'Systems of lateral sounds and perception', *Proceedings of the 6th International Congress of Phonetic Sciences* (1967), Prague.
- Holmes, Janet (1984) 'Hedging your bets and sitting on the fence: some evidence for hedges as support structures', *Te Reo*, 27: 47–62.
- Hopper, Paul J. (1997) 'Dispersed verbal predicates in vernacular written narrative', in A. Kamio (ed.), 1–17.
- Hopper, Paul J. and Traugott, Elizabeth C. (1993) *Grammaticalization*, Cambridge: Cambridge University Press.
- Jackson, Kenneth (1968) 'The breaking of original long *é* in Scottish Gaelic', in J. Carney and D. Green (eds.) *Celtic Studies: Essays in Memory of Angus Matheson 1912–1962*, 65–71, London: Routledge.

- Jang, Shyue-Chian (1998) 'Dimensions of spoken and written Taiwanese: A corpus-based study', PhD dissertation, University of Hawaii, UMI Number 9829561.
- Jucker, A.H. (1992) *Social stylistics: syntactic variation in British newspapers*, Berlin: Mouton de Gruyter.
- Kamio, Akio (1997) 'Preface', in A. Kamio (ed.), ix–xi.
- Kamio, Akio (ed.) (1997) *Directions in Functional Linguistics*, John Benjamin: Amsterdam/ Philadelphia.
- Kim, Yong-Jin and Biber, D. (1994) 'A corpus-based analysis of register variation in Korean', in D. Biber and E. Finegan (eds.), 157–81.
- Kirk, John M. (1997) 'Subordinate clauses in English', in *Journal of English Linguistics*, 25(4): 349–64.
- Konstantopoulos, Stasinou (1998) *A Morphological Analyser for Scottish Gaelic*, unpublished MSc dissertation, University of Edinburgh.
- Kroll, B. (1977) 'Combining ideas in written and spoken English: A look at subordination and coordination', in E. Ochs Keenan and T.L. Bennet (eds.) *Discourse across Time and Space: Southern California Occasional Papers in Linguistics*, 5: 69–108, Los Angeles: University of Southern California.
- Lakoff, R. (1984) 'The pragmatics of subordination', *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*: 481–92, Berkeley: University of California.
- Lamb, William (1999) 'A diachronic account of Gaelic News-speak: The development and expansion of a register', *Scottish Gaelic Studies*, XIX: 141–71.
- Lamb, William (2002) *Scottish Gaelic*, Second Edition, Languages of the World/Materials, München: Lincom Europa.
- Lamb, William (2005) 'Toward a corpus of spontaneous spoken Scottish Gaelic', in J.M. Kirk and D.P. Ó Baoill (eds.) *Legislation, Literature and Sociolinguistics: Northern Ireland, the Republic of Ireland, and Scotland*, 126–37, Belfast: Cló Ollscoil na Banríona.
- Lamb, William (2006) 'The Noun Phrase in Scottish Gaelic Speech and Writing', in W. McLeod, J.E. Fraser and A. Gunderloch (eds.) *Cànan & Cultar / Language & Culture: Rannsachadh na Gàidhlig* 3, 153–78, Edinburgh: Dunedin Academic Press.
- Lambrecht, Knud (1994) *Information Structure and Sentence Form*, Cambridge: Cambridge University Press.
- Leech, G. (1991) 'The state of the art in corpus linguistics', in K. Aijmer and B. Altenberg (eds.) *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, 8–29, London: Longman.
- Leech, G., Myers, G., and J. Thomas (eds.) (1995) *Spoken English on Computer: Transcription, Mark-up and Application*, London: Longman.
- Lehmann, Christian (1985) 'Grammaticalization: synchronic variation and diachronic change', *Lingua e Stile*, 20: 303–18.
- Lyons, John (1999) 'Sentences, clauses, statements, and propositions' in P. Collins and D. Lee (eds.), 149–76
- Macaulay, D. (1966) 'Palatalization of labials in Scottish Gaelic and some related problems in phonology', *Scottish Gaelic Studies*, 11: 72–84.
- Macaulay, D. (1979) 'Some functional and distributional aspects of intonation in Scottish Gaelic: A preliminary study of tones', *Occasional Papers in Linguistics and Language Learning*, 6: 27–38.
- Macaulay, D. (1982a) 'Aspects of register range and choice in Scottish Gaelic', *International Journal of the Sociology of Language* 35: 25–48.

- Macaulay, D. (1982b) 'Borrow, calque, and switch', *Current Issues in Linguistic Theory* 15: 203–37.
- Macaulay, D. (1986) 'New Gaelic', *Scottish Language* 5: 120–5.
- Macaulay, D. (1992) 'The Scottish Gaelic Language', in D. Macaulay (ed.) *The Celtic Languages*, 137–248, Cambridge: Cambridge University Press.
- Macaulay, D. (1996) 'Some thoughts on time, tense, and mode, and on aspect in Scottish Gaelic', *Scottish Gaelic Studies* 14: 193–204.
- Macaulay, R.K.S. (1991) *Locating Dialect in Discourse*, Oxford: Oxford University Press.
- MacDonald, John A. (1999) *Òrain Dhòmhnaill Ailein Dhòmhnaill na Bainich*: The Songs of Donald Allan MacDonald 1906–92, An Gearasdan.
- MacDonald, S. (1987) 'Social and Linguistic Identity in Staffin, Isle of Skye', unpublished D.Phil thesis, Oxford University.
- MacDonald, William (2000) Personal communication. [Mr. MacDonald is the headmaster of Uist's secondary school.]
- MacInnes, J. (1992) 'The Scottish Gaelic Language', in G. Price (ed.) *The Celtic Connection*, 101–30, Gerrards Cross: Colin Smythe.
- MacInnes, Paul (1994) 'Gaelic Drama', in D. Thomson (ed.) (1994a).
- MacKinnon, Kenneth (1974) *The Lion's Tongue*, Inbhir Nis: Club Leabhar.
- MacKinnon, Kenneth (1977) *Language, Education, and Social Processes in a Gaelic Community*, London: Routledge and Kegan Paul.
- MacKinnon, Kenneth (1994) 'Gaelic language use in the Western Isles', in A. Fenton and D.A. MacDonald (eds.) *Studies in Scots and Gaelic: Proceedings of the Third International Conference on the Languages of Scotland*, Edinburgh: Canongate Academic and The Linguistic Survey of Scotland.
- MacKinnon, Kenneth (1998) 'Gaelic in family, work and community domains: Euromosaic Project 1994/95', *Scottish Language*, 17: 55–69.
- MacKinnon, Kenneth (2000) 'Scottish Gaelic' in G. Price (ed.) *Languages in Britain and Ireland*, 44–57, Oxford: Blackwell Publishers Ltd.
- MacPherson, John (1945) *The North Uist Dialect*, unpublished PhD thesis, University of Edinburgh.
- MacThòmais, R. (1976) *Bidh-Eòlas: A' Chealla, Ginntinneachd, is Mean-fhàs*, Gairm: Glaschu.
- Mathieson, C. and Thompson, S. (1988) 'The structure of discourse and 'subordination'', in J. Haiman and S. Thompson (eds.) *Clause Combining in Grammar and Discourse*, 275–329, Amsterdam and Philadelphia: John Benjamins.
- McCarthy, M. (1998) *Spoken Language and Applied Linguistics*, Cambridge: Cambridge University Press.
- McEnery, T. and Wilson, A. (1996) *Corpus Linguistics*, Edinburgh: Edinburgh University Press.
- McEnery, T. and Ostler, N. (2000) 'A new agenda for corpus linguistics: Working with all of the world's languages', *Literary and Linguistic Computing*, 15(1): 403–19.
- McLeod, Wilson (2000) 'Official Gaelic: Problems in the translation of public documents', *Scottish Language*, 19: 100–16.
- McNeir, Clive Leo (ed.) (2001) *Faclair na Pàrlamaid: Dictionary of Terms*, The Scottish Parliament: Edinburgh.

- Meek, D. (1990) 'Language and style in the Scottish Gaelic Bible (1767–1807)', *Scottish Language*, 9: 1–16.
- Miller, Jim (2000) 'Focus in the languages of Europe', *Eurotype*.
- Miller, Jim, and Weinert, Regina (1998) *Spontaneous Spoken Language*, Oxford: Clarendon.
- Mithun, Marianne (1985) 'How to avoid subordination', *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*: 493–509.
- Mithun, Marianne (1988) 'The grammaticization of coordination', in Haiman, J. and S. Thompson (eds.) *Clause Combining in Grammar and Discourse*, 331–57, Amsterdam and Philadelphia: John Benjamins.
- Murchison, T.M. (1994) 'Crofting System', in D. Thomson (ed.) (1994).
- Nelson, Gerald (1997) 'Cleft constructions in spoken and written English', *Journal of English Linguistics*, 25(4): 340–8.
- Nicholson, Alexander (1936) *Modern Gaelic: A Basic Grammar*, Glasgow: Alex Maclaren and Sons.
- Ní Laoire, Siobhán (1988) 'Preliminaries to a study of register in Modern Irish', *Language, Culture and Curriculum*, 1(3): 289–301.
- Ní Laoire, Siobhán (1993) 'Traditions of spoken language study in Ireland', *Irish Review*, 14: 65–73.
- Nivet, Marie-Christine (c. 1989) *Syntax of French Conversation: A Comparative Study of Two French Speakers*, PhD dissertation, Ann Arbor: UMI.
- Oakes, Michael P. (1998) *Statistics for Corpus Linguistics*, Edinburgh: Edinburgh University Press
- Ó Dochartaigh, C. (ed.) (1994–7) *Survey of the Gaelic Dialects of Scotland: Questionnaire Materials Collected for the Linguistic Survey of Scotland*, 5 vols., Dublin: Dublin Institute for Advanced Studies.
- O'Donnell, Roy C. (1974) 'Syntactic differences between speech and writing', *American Speech*, 49: 102–10.
- Ochs, E. (1979) 'Planned and unplanned discourse', in T. Givón (ed.), 51–80, London: Academic Press.
- Oftedal, Magne (1956) 'The Gaelic of Leurbost, Isle of Lewis', *Norsk Tidsskrift for Sprogvidenskap, Suppl. Bind IV: A Linguistic Survey of the Gaelic Dialects of Scotland*, vol. III., Oslo: H. Aschehoug and Co.
- Ó Maolalaigh, Roibeard (1996) *Scottish Gaelic in 3 Months*, with Iain MacAonghais (consultant), Hugo.
- Ó Maolalaigh, Roibeard (1997) 'Preaspiration in Scottish Gaelic dialects', unpublished manuscript, University of Edinburgh.
- Ó Maolalaigh, Roibeard (1999) 'Transition zones, hyperdialectisms and historical change: the case of final unstressed *-igh/-ich* and *-idh* in Scottish Gaelic', in *Scottish Gaelic Studies*, XIX: 195–233.
- Ó Murchú, M. (1989) *East Perthshire Gaelic*, Dublin: Dublin Institute for Advanced Studies
- Ó Siadhail, M. (1984) 'Agus (Is)/And: A shared syntactic feature', *Celtica*, 26: 125–37.
- Ó Siadhail, M. (1989) *Modern Irish: grammatical structure and dialectal variation*, Cambridge: Cambridge University Press.
- Payne, T.E. (1997) *Describing Morphosyntax: A Guide for Field Linguists*, Cambridge: Cambridge University Press.
- Poole, M.E., and Field, T.W. (1976) 'A comparison of oral and written code elaboration', *Language and Speech*, 19: 305–11.

- Prince, Ellen F. (1997) 'On the functions of Left-Dislocation in English discourse', in A. Kamio (ed.), 117–43.
- Ramchand, Gillian C. (1995) *Aspect and Predication: The Semantics of Argument Structure*, Oxford: Oxford University Press.
- Romaine, S. (1994) 'On the creation and expansion of registers: Sports reporting in Tok Pisin', in D. Biber and E. Finegan (eds.) *Sociolinguistics Perspectives on Register*, 59–82, New York: Oxford University Press.
- Sapir, Edward (1974) 'Language', in Ben G. Blount (ed.) *Language, Culture and Society*, 46–66, Cambridge: Withrop Publishers, Inc., reprinted from *The Encyclopedia of the Social Sciences*, Seligman and Johnson (ed.), vol. 9, 1933, 155–69, Macmillan Company.
- Shaw, John (1999) 'The ethnography of speaking and verbal taxonomies: Some applications to Gaelic', in W. Gillies, R. Black, and R. Mullally (eds.) *Celtic Connections: Proceedings of the 10th International Congress of Celtic Studies*, 309–23.
- Schiffrin, Deborah (ed.) (1984) *Meaning, Form, and Use in Context: Linguistic Applications*, Washington: Georgetown University Press.
- Schiffrin, Deborah (1987) *Discourse Markers*, Cambridge: Cambridge University Press.
- Svartvik, Jan (ed.) (1992) *Directions in Corpus Linguistics: Proceeding of Nobel Symposium 82*, Stockholm, 4–8 August 1991, Mouton de Gruyter: Berlin and New York.
- Tannen, Deborah (ed.) (1982a) 'The oral/literate continuum in discourse', in D. Tannen (ed.) *Spoken and Written Language: Exploring Orality and Literacy*, 1–16, New Jersey: Ablex.
- Tannen, Deborah (1982b) 'Oral and literate strategies in spoken and written narratives', *Language*, 58: 1–21.
- Ternes, E. (1973) *The Phonemic Analysis of Scottish Gaelic*, Hamburg: Helmut Buske Verlag
- Ternes, E. (1994) 'Gaelic: phonemic structure', in D. Thomson (ed.) (1994a), 101–4.
- Thomson, D. (1971) 'Scottish Gaelic literature: The contemporary situation', in J.E. Caerwyn Williams (ed.) *Literature in Celtic Countries*, 79–96, Cardiff: University of Wales Press.
- Thomson, D. (1979) 'Gaelic: Its range of uses', in A.J. Aitken and T. Mc Arthur (eds.) *Languages of Scotland*, Edinburgh: W and R Chambers.
- Thomson, D. (1989) 'Gaelic publishing', *Scottish Language*, 8: 34–41.
- Thomson, D. (1990) 'Gaelic in Scotland: Assessment and prognosis', in E. Haugen, J. Derrick McClure and D. Thomson (eds.) *Minority Languages Today*, 10–20, Edinburgh: Edinburgh University Press.
- Thomson, D. (1994) 'Attitudes to linguistic change in Gaelic Scotland', in M. Parry, W. Davies and R. Temple (eds.) *The Changing Voices of Europe*, 227–35, Cardiff: University of Wales Press.
- Thomson, D. (ed.) (1994a) *The Companion to Gaelic Scotland*, Glasgow: Gairm Publications.
- Thompson, Sandra A. (1984) "'Subordination" in formal and informal discourse', in D. Schiffrin (ed.), 85–94.
- Thompson, Sandra A. (1988) 'A Discourse Approach to the Cross-Linguistic Category 'Adjective'', in *Explanations for Language Universals*, John Hawkins (ed.), 167–85 Oxford: Basil Blackwell.

- Thompson, Sandra A. (1997) 'Discourse motivations for the core-oblique distinction as a language universal', in A. Kamio (ed.), 59–82.
- Thurneyson, Rudolf (1993) *A Grammar of Old Irish*, Dublin: Dublin Institute for Advanced Studies.
- Van Valin, R. (1985) 'A typology of syntactic relations in clause linkage', in Brugman *et al.*, 542–57.
- Van Valin, R. and LaPolla, R. (1997) *Syntax: Structure, Meaning, and Function*, Cambridge: Cambridge University Press.
- Watson, W.J. (1927) 'Some interactions between Gaelic and English', in *Transactions of the Gaelic Society of Inverness*, XXXIII: 310–26.
- Withers, C. (1984) *Gaelic in Scotland, 1698–1981*, Edinburgh: John Donald.

Index

A

- abstractness, level of, 26, 33, 34
- academic discourse, 18, 26
- academic prose, 21, 33, 50, 52
 - definition, 54
 - situational description and examples, 60-1
 - texts for study, 54
- address forms, 35
- Adger, David, 249, 297
- adjectives, 65, 116-20
 - attributive, 25, 26, 33, 65, 116, 117, 219-20
 - case-marking, 156-7, 211-12
 - comparative and superlative, 65, 116, 118-19, 221-2
 - demonstrative, 65, 116, 117-18, 119, 149
 - elaboration by, 25
 - nonconcordance in, 156
 - postnominal, 219-20
 - predicative, 65, 116, 118, 219, 220, 246
 - prenominal, 205, 220
 - quantitative, 221
 - summary, 219-22
- adverbial clauses, 29, 257-62
 - bound and free, 259
 - concessive, 259
 - conditional, 260-1
 - finite, 27
 - location, 261-2
 - manner, 261
 - position of, 27
 - reason, 96, 258-9
 - temporal, 84
 - types of verbal complement, 258
- adverbials, 66, 68
 - amplifiers, 66, 137-8, 140, 224
 - downtoners, 66, 138, 139, 224
 - emphatic, 66, 138, 139, 140, 224
 - fronting, 81-3, 100
 - hedges, 66, 138-9, 140, 224
 - peripheral modifiers, 81-2
 - place, 66
 - placement of, 64, 81-4
 - time, 66
- adverbs, 66, 135-7
 - degree, 224
 - direction and location, 224
 - elaboration by, 25
 - grammatical decisions, 68-9

- manner, 223
- place, 135-6, 178
- of stance and degree, 137-40
- summary, 222-4
- time, 33, 34, 135, 137, 223
- aesthetic orientation, 23-4
- affixation
 - emphatic, 265
 - impersonal, 233
 - personal in finite and non-finite verbs, 232-3
- age factors, 22, 25, 36, 43, 44, 46, 54, 55
- agreement, 30
- Akinassio, F.Niyi, 26
- amplitude, 29
- anaphora, 119, 132, 206
- ANOVA (analysis of variance), 18, 70
- anthropology, and linguistics, 23
- Applecross, 199
- apposition, 81, 120
- appositives, 65, 120-2, 184, 210
- archaic forms, 40
- archive, 52
- argumentation, 34
- Arran, 201
- article, definite, 208-9
- aspect, 102-5
 - perfect, 65, 102-3, 237
 - periphrastic, 238
 - pluperfect, 237
 - progressive, 25, 65, 103-5, 177, 237
 - prospective, 65, 105, 237
 - summary, 237-8
- aspiration *see* lenition
- asyndesis, 26, 86, 90, 97, 193
- authorship, 22
- auxiliary verbs, 25, 234, 237, 238-9, 243, 251

B

- Beaman, Karen, 28, 93
- beliefs, 44
- Benbecula, 40, 42, 43, 84, 275
- Biber, Douglas, 17, 19, 20, 22, 23, 24, 28, 29, 31, 37, 52, 53, 55, 63, 73, 74, 76, 77, 81, 90, 91, 92, 94, 95, 96-7, 101, 103, 105, 106, 107, 109, 111, 112, 116, 119, 120-1, 123, 128, 131, 132, 133, 135, 136, 137, 138, 139, 140-1, 142, 145, 146, 147, 148, 150, 157, 172, 175, 179, 181, 188, 194, 196, 297, 299, 300, 301, 305
- dimensions of variation, 33-4, 191-3
- on English variation, 18, 33, 190
- multidimensional analysis, 32-4
- text types, 34
- universal dimensions of register variation, 33-4, 38

Biblical language, 36, 50
 bilingualism, 17, 47-50
 Black, Ronald, 49
 Blankenship, Jane, 24
 body posture, 29
 Borgstrom C.H.G., 40, 198, 208, 297, 298
 borrowing from English, 35, 36, 47, 50, 66, 143-5, 188, 190, 199
 assimilated nouns, 143-4, 273
 in individual registers, 175, 176, 178, 180, 181, 182, 183, 185
 nativisation, 68
 partly assimilated, 273
 unassimilated non-nouns, 145, 149
 unassimilated nouns, 144, 273
 see also calques
 breaking, of Old Irish long é, 40
 British National Corpus, 30
 broadcasts, Gaelic, 33, 48, 54, 58
 Brown Corpus, 29
 Bruford, Alan, 59, 298

C
 Calder, G., 202, 203, 206, 222, 223, 224, 298
 calques, 51, 274
 Campbell, John Francis, 23
 Campbell, John Lorne, 35
 case, 65
 grammatical decisions, 68-9
 summary, 209-12
 see also dative case; genitive case; nominative case
 case-marking, 19, 150-7
 adjectives, 156-7, 211-12
 concordance and nonconcordance, 150-7
 loss of genitive, 36
 nouns, 210-12
 causatives, 241-2
cèilidh, 35, 43
 Celtic languages, 39
 Census(1981), 46
 (1991), 42, 44, 49, 51
 (2001), 42
 Chafe, Wallace, 17, 18, 23-4, 27, 28, 29, 30, 33, 37, 38, 52, 63, 81, 84, 92, 93, 94, 107,
 109, 116, 126, 131, 132, 138, 141, 145-6, 150, 157, 182, 186, 189, 191, 193, 194, 259,
 270, 298
 effect of mode on linguistic divergence, 26
 Chi Square, 25
 children
 number of Gaelic-speaking, 42, 44, 45, 46
 speaking only English, 44
 television programmes for, 48, 49
 see also early child discourse

- Chomsky, Noam, 19
- clause clusters, 30
- clause combinations, 251-64
- clause types, 64, 73-101
- clauses
 - defined, 31, 66-7, 252
 - dependent, 25
 - different kinds, 25, 32
 - finite, 27, 31
 - juxtaposition of related, 28
 - layered structure, 252
 - negative, 65, 111-12, 148
 - nuclear and peripheral, 27
 - peripheral modifying, 27
 - relationships between, 27
 - structure of, 30, 252
 - temporal, 34
 - transitive, 171
 - as unit of analysis, 30-1, 66-7
 - see also* adverbial clauses; coordinated clauses; cosubordinate clauses; imperatives; interrogative clause; main/matrix clauses; passive clauses; relative clauses; small clauses; subordinate clauses
- cleft constructions, 27, 68, 267-8
- clefts, 27, 73-7
 - deictic, 64, 269
 - demonstrative, 73, 75, 76, 100, 172, 173
 - highlighting by, 266-70
 - identificational, 65, 115-16, 245, 267
 - nominal-type, 267
 - non-nominal type, 268-9
 - overall results, 74
 - 's *ann*-type, 64, 73, 74-5, 76, 100
 - 's *e*-type, 64, 73, 75, 76, 100, 172, 176
 - WH *see* pseudo-clefts
- code-switching, 35, 36, 47, 50, 55, 219, 274-5
 - gender and, 46
- cohesion, 26, 28, 46, 116, 119, 178, 181, 183
- colloquial speech, 19, 23, 27, 28, 29
- communication
 - alternative form to speech, 23, 29
 - face-to-face, 29, 57
- community, 42-4
- compass directions, 224
- competence, 19
- complement clauses, 27, 29, 64, 93-5, 101, 256-7
 - avoidance in speech, 98
 - distinguishing news scripts, 94
 - finite or non-finite, 27, 28, 257
 - indirect questions, 257

- complex NPs (noun phrases), 65, 122-3, 157, 158-9, 195
 - single adjective, 159-61
 - single prepositional, 161-4
 - single relative clauses, 164-6
 - two or more adjectives, 167-9, 195
 - two or more genitives, 166
 - two or more prepositional phrases, 170
 - two or more relative clauses, 170-1
- complexity
 - of adjectives in speech, 118
 - clausal, 73, 88
 - linguistic, 28
 - noun phrase, 71, 93, 122, 131, 150-71, 174, 181, 292
 - relative clauses in formal prose, 164
 - scoring of, 122
 - in spoken language, 25, 28-9
 - and subordination, 29
 - syntactic, 17
 - in writing, 28-9, 164
- concessive clauses, 64, 95-6, 259
- conditional clauses, 64, 96, 260
 - counterfactuals, 260-1
 - hypothetical, 260
 - negative, 261
- conditional-future, 236
- conjunction, 254-5
- conservativeness, 187
- consonant mutation, 199-201
- consonants, 199
- Constantinidou, Evi, 46
- constituent order, and grammatical relations, 248-51
- construction-type, 17
 - literate, 18
 - regularity of appearance in speech or writing, 24
 - variation, 20
- constructs, problems of definition, 27-8
- context
 - changes, 17, 37
 - factors in individual register, 32
 - situational, 19, 31
 - see also* social context
- contextual variation, and linguistic form, 21, 31-2, 37, 188-90
- contractions, 17, 27, 34
- convention, 32
- conversation, 17, 21, 52, 194
 - in corpus, 53, 55
 - difference from other forms of language, 27
 - English borrowing in, 145
 - as fundamental, 23-4, 29

- observations on, 186
- profile of register, 172-5
- properties shared with media language, 194
- scope in, 76-7
- situational description, 55-6
- subordinate clauses in, 18
- texts for study, 53
- see also* Spontaneous Spoken Language (SSL)
- coordinate clauses, 64, 84-6, 100
 - strings of, 27
- coordination, 28, 252, 254-6
 - conjunction, 254-5
 - disjunction, 255
 - exclusion, 255-6
 - of intra-clause constituents, 27
 - in traditional narrative, 179
- coordinators, initial, 26
- copular constructions, identificational clefts, 65, 245
- copular verbs, 65, 114-15
 - past-conditional, 245, 246
 - plus adjectives, 221-2
 - present-future, 245
- core of clause, 31, 252
- corpus
 - of academic discourse, 26
 - coding of, 52, 63
 - compared with archive, 52
 - computerised, 18, 21, 31, 32, 37, 196
 - definition, 22, 31
 - see also* named corpora
- Cos, Richard, 50, 68, 299
- cosubordinate clauses, 31, 87-8, 100
 - with *agus*, 64
- cosubordination, 250, 252, 263-4
 - definition, 264
 - participle type, 263
 - reason type, 263
 - relative clause type, 264
 - temporal type, 264
- Cram, David, 20, 229, 274, 299
- crofting, 42-3, 45-6, 47, 55
- Crowdy, S., 55, 299
- culture, breakdown of traditional, 43

D

- Danielewicz, Jane, 27, 28, 30, 84, 131, 132, 138, 146, 186, 189, 298
- dative case, 209, 210
 - concordance data, 152-3
- declaratives, 238, 239

- definiteness/referentiality, 208-9
- definition
 - inconsistency of, 29
 - problem of construct, 27-8
- deixis, 66, 73, 76, 77, 247, 264, 269
- demography, 42
- demonstratives, 66, 133, 215
- deontic modality, 238-9
- depopulation, 43
- descriptive grammar, 21, 64, 197-286
- detachment, 19, 26, 29, 37, 77-81, 191, 194
 - centre, 271
 - left, 25, 29, 64, 77, 78, 79-80, 98, 99, 100, 195, 265, 270
 - right, 64, 78, 80, 100, 271
- determiners, kinds of, 26
- diachronic studies, 22, 32, 35, 243, 302
- dialectology, 35, 36, 38, 39-42
- dialects, 22
 - 'central' and 'peripheral' group, 40
 - endangered, 35
 - features of Sc Gaelic, 40
 - prestige forms, 36
- dictionaries
 - of broadcasting and parliamentary terms, 48
 - Cox's, 68
 - Dwelly's, 68
 - use in determining grammatical categories, 68
- diglossia, limited, 36
- diminutives, 228
- diphthongisation, 40
- diphthongs, 199
- discourse
 - early child compared with adult spontaneous speech, 25
 - information structure and, 64
 - institutionalised higher, 19
 - organisation, 30, 32
 - planned, 25
- discourse analysis, 20
 - of a single language variety, 27
- discourse freedom, 71, 149, 189
- discourse phenomena, 264-73
 - highlighting of constituents, 266-71
 - interjections, 272-3
 - parenthetical statements, 272
 - particles, 271
 - referential devices, 264-5
- disfluencies, 23, 26, 27, 68
- disjunction, 255

- domain, 47
 - domestic, 17, 47, 55
 - ritual, 17, 26-7
 - status of, 37
- Dorian, Nancy, 19, 20, 36, 37, 38, 44, 47, 145, 198, 199, 266, 299
- drama, Gaelic, 50
- Dressler, Wolfgang, 17, 21, 36-7
- Dunnett's T3 pairwise comparisons test, 70

E

- e-mail, 49
- early child discourse, compared with adult spontaneous speech, 25
- East Sutherland Gaelic (ESG), 36, 44, 46, 199
- economy, 42-4, 50
- education
 - Gaelic-medium, 49-50, 195
 - level of, 25, 55
- electronic media, 60
- ellipsis, 27, 87, 177, 188
- emphatics *see* adverbials, emphatic
- endangered languages, 21, 37
- English, 18, 19, 20, 27, 30, 32, 34, 38, 42, 43, 60, 62, 63, 72, 73, 74, 77, 79, 100, 103, 105, 107, 119, 122, 130, 131, 157, 158, 160, 186, 191, 193, 194, 195, 196, 205, 215, 216, 217, 219, 220, 224, 231, 234, 239, 242, 243, 244, 249, 255, 256, 259, 264, 266, 268, 269, 270-1, 274, 292, 293, 294, 297, 298, 299, 300, 301, 302, 304, 305, 306
 - acquisition of, 44
 - influence of, 273-5
 - lexical obscenity in, 272-3
 - passivisation in, 34
 - register differences, 17
 - see also* borrowing from English; language usage, Gaelic-English; Scottish English
- epistemic modality, 239
- ethnology of communication acts, 55
- ethnography, 42
- Euromosaic Project, 44, 45
- exclamations, 272-3
- exclusion, 255-6
- existentials, 247-8

F

- facial expressions, 29
- factor analysis, 33, 37, 172, 188
- false starts, 25, 68, 79
- familiarity, 35
- features
 - contextual, 21, 188-90
 - definition, 22
 - linguistic, 21

- relative frequencies of core linguistic, 32, 70, 186
- selection for register study, 64-6
- Ferguson, C., 22, 31, 32, 52, 300
- fiction, 21, 34, 49, 50, 52
 - adventure, 33
 - mystery, 33
 - profile of register, 180-1
 - romantic, 33
 - situational description and examples, 61
 - texts for study, 54
- fishing, 43, 47
- fluency, 22
 - semi-fluent speakers, 44, 47
 - see also* disfluencies
- focus constructions, 19, 26, 30
- Foley, W., 29, 251, 300
- folklore, 23, 35, 50, 59
- formal discourse, vs. colloquial discourse, 28, 35
- formal prose, 17, 60
 - characteristics of, 186-7
 - compared with spontaneous conversation for complexity and function, 150-71
 - profile of register, 181-2
- formality, 19, 47, 72, 85, 143, 164, 175, 182, 184, 190, 274
 - concentration on, 19, 20, 36
 - and conservative nominal forms, 150
 - mode and, 27, 31
 - nonconcordance as measure of, 125
- formulae *see* register markers
- fragmentation, 26, 28, 30, 34, 37, 77, 161, 167, 171, 173, 174, 191, 194
- French, 17, 20, 48, 135, 202, 214, 244, 304
 - conversation study, 52
 - transcription of speech, 30
- fronting, 240, 264, 270
 - adverbial, 81-3, 100, 181, 187, 189, 194
 - nominals, 270
 - subordinate clauses, 83, 84
 - see also* highlighting of constituents
- function, 18, 21, 22, 32, 33-4
 - communicative sharing, 186
- functional grammar, 24
- fusion, 202
- future tense, conditional-future, 236

G

- Gaelic, *see also* Irish; Manx Gaelic
- Gaelic (Scottish)
 - broadcasting, 33, 48
 - decline of, 45, 46
 - dialects, 40

- history of, 39-42
- intergenerational use, 44-51
- literacy, 49-50
- maintenance of, 46
- origins, 39
- speakers' assessment of, 44
- transmission of, 42, 44, 50
- Gàidhealtachd, 42
 - map, 39
- gender, 28
 - defective, 206
 - and Gaelic-English usage, 46
- General Linear Model (GLM), ANOVA, 70
- generative grammar, 19, 20, 24
- genitive case, 26, 126-7, 209-10, 211
 - concordance and nonconcordance, 153-5
 - loss of marking, 36
- genres
 - traditional, 36
 - use of term, 22, 31
 - written range, 49
- German, 23, 29, 122, 135, 158, 214
- gerunds, 25, 244-5
- gestures, 23, 29
- glossing, 121
- Goidelic, 20, 21, 38, 39, 40, 150, 249, 271, 297
 - greater Goidelic continuum, 40
- government, 30
- grammar
 - difficulties in resolving categories, 68
 - prescriptive, 19, 24, 195
 - see also* descriptive grammar; functional grammar; generative grammar; Role and Reference Grammar
- grammatical relations, and constituent order, 248-51
- Greenbaum, Sydney, 18, 27, 28, 30, 38, 52, 63, 84, 86, 94, 177, 300, 301
- greeting language, high register, 36
- Grimm, Jacob, 23

H

- Halliday, M.A.K., 23, 28, 29, 37
- have, to* expressions, 213
- head/dependent marking, 204-5
- hedges, 26, 27, 34, 66, 138-9, 140, 175, 176, 178, 179, 183, 187, 191, 193, 224, 239, 271, 301
- highlighting of constituents, 266-71
 - emphatic suffix, 266
 - word order, 266-71
- honorifics, 135, 214
- Human Communication Research Centre, 16

hypercorrection, 152
hypotaxis, 30

I

Iain Óg Ìle 23
Icelandic, 201
idiolect, 24, 25, 214, 273
illocutionary force, 229, 238, 239-40, 253, 264
imperatives, 65, 112, 113-14, 148, 238
 direct and indirect, 240-1
 negative, 241
imperfect aspect, 24
impersonals, 110-11
 synthetic, 65
 as valence decreasing expressions, 242, 244
imprecations, 272-3
incomers, 43, 50
indefiniteness, 208
 marking of nouns, 65, 127-8, 148
 pronouns, 217-18
Indo-European, 39, 231, 251
inferential statistics, 18, 25, 32
infinitive constructions, 25
informal speech *see* conversation
informality, 17, 18, 19, 26, 30, 77, 158, 205
information structure, 21, 53, 64, 71, 73-101, 102, 194, 271
information-based registers, 33, 34, 63
informationality, 71, 72, 146, 148, 188, 189, 191
ingression, 272
initial mutation *see* lenition
integration, 26, 28, 30, 37, 181, 191, 194
interactiveness, 33, 34, 148, 187, 189
interjections, 36, 272-3
International Corpus of English, 27, 76
interrogatives, 65, 112-13, 148, 238
 pronouns, 216-17
intonation, 28, 32, 55, 259, 302
 utterance-final, 30
intuitions as a data source, 24, 28
involvement, 26, 33, 37, 96, 139, 141, 174, 175, 176, 179, 184, 188, 191, 194, 298
Irish, 19, 20, 39, 40, 61, 104, 165, 166, 208, 218, 219, 221, 224, 266, 270, 271, 304, 306
 conversation studies, 35
 Old Irish, 40, 202
Islay, 199
isoglosses, 40

J

Jackson, Kenneth, 40
Jang, Shyue-Chian, 20, 32, 52, 54, 62, 63, 96, 97, 103, 112, 114, 115, 116, 123, 131,

- 135, 136, 137, 138, 139, 141, 142, 146, 192, 193, 301
- text types, 34
- Jones, Daniel, 23
- Jucker, A.H., 157, 301
- juncture, 252
 - clausal, 253
 - core, 253
 - nuclear, 252
- juxtaposition, 26, 28, 73, 83, 86, 92, 93, 97, 101, 186, 188, 191, 193, 194, 195, 245, 254
 - flexibility on, 99, 100
 - versus subordination, 97-100, 101, 195

K

- Kim, Yong-Jin, 192-3, 300
- Kintyre, 201
- knowledge
 - shared, 55, 59
 - specialised, 60
- Korean, 32
- Kruskal-Wallis test, 70

L

- Lackoff, R., 264
- Lambrecht, Knud, 29, 77
- Lancaster Oslo-Bergen Corpus of British English, 18, 52
- language change, internal, 37
- language choice *see* language usage, Gaelic-English
- language contact, 19, 44
- language death, 17, 36-7, 38, 195, 299
- language loyalty, gender and, 46
- Language Maintenance and Viability Survey, 44, 45
- language usage, Gaelic-English, 44-51
- language/linguistic variety, 22
 - discourse properties of a single, 27
- languages, idiosyncratic oral/literate dimensions, 34
- Latinate paradigms, 21
- lectures, 24, 30
- Leech, G., 52, 297, 298, 299, 300, 301, 302
- lenition, 199-200, 205, 208, 210, 212, 218, 220, 221
- letter writing, 49
- Levene's test, 70
- Leverhulme (?Leverhume) Corpus, 27, 52
- Lewis, Isle of, 40, 42, 43, 160, 201, 280, 304
- lexical choice, 32
- lexical classes, 66, 137-46
- lexical obscenity, 272-3
- lexical range, 27, 148, 150, 162, 171, 176, 195
- lexical specificity, 66, 146-8
- lexical variety *see* type-token ratio

lexicalisation, 26
 lexicography, 196
 lexicon, 17, 32, 71, 102-49, 148, 181, 190, 231, 273
 density of written, 28
 Gaelic-based for radio, 58
 lexis, categories of, 19
LinguaStat, 68, 69, 151, 157, 196
 linguistic description, 32
 linguistic form, and contextual variation, 21, 32, 37, 188-90
 linguistic typology, 24
 linguistics, and anthropology, 23
 literacy, 23
 Gaelic, 49-50
 literature, Gaelic, 49-50
 loan vocabulary, 47-8
 locatives, 246-7
 Lochalsh, 45
 London-Lund Corpus of Spoken English, 18, 52
 Lyons, John, 30

M

Macaulay, Donald, 20, 35, 36, 47, 229, 231, 233
 tenses paradigm, 234-6
 Macaulay, R.K.S., 68
 McEnery, T., 31, 196, 297, 303
 MacInnes, John, 42, 289, 303
 MacKinnon, Kenneth, 42, 43, 44, 45-6, 165, 303
 Maclean, Calum, 35
 McLeod, Wilson, 49
 'magnasyntax', 19
 magnavocabulary, 47, 48
 main/matrix clauses, 64, 83, 88-90, 100
 Man, Isle of, 40
 Manx Gaelic, 39
 Map Task, 29
 markers, in register analysis, 31-2, 34
 meaning
 conveyed by syntax and spoken rhythm, 26
 lexicalisation in writing, 26
 media language, 32, 48-9, 194
 properties shared with conversation, 194
 Meek, Donald, 19, 36, 50, 303
 migration, 46
 Miller, Jim, 17, 18, 19, 23, 24, 27, 29, 30, 31, 38, 52, 53, 63, 73, 76, 77, 79, 84, 86, 96,
 116, 122, 123, 130, 131, 150, 157, 158, 159, 179, 186, 194, 195, 231, 266, 270, 303
 on relative clauses, 92
 on subordinate clauses, 90
 Milroy, Leslie, 16
 minority languages, 17

- modal/perfective auxiliaries, 25, 238-9
- mode, 17, 18, 19, 24, 26, 28, 29, 34, 105, 127, 133, 163, 190, 230, 233, 234, 235, 260, 297, 302
 - and clause combination variation, 27
 - definition, 22
 - direct interrogative clause, 65, 112-13
 - and formality, 27, 31
 - imperative verbs, 65, 113-14
 - irrealis, 238
 - linguistic differences due to, 24-5, 26, 75
 - problems with definition, 27-8, 238
 - realis, 238
 - and register conflated, 28
 - summary, 238-41
 - variation within, 31
 - see also* deontic modality; epistemic modality; illocutionary force; speech and writing
- modification types, 32
- monoglots
 - English, 47, 50
 - Gaelic, 44
- monologue, 27, 72, 74, 85, 100, 134, 189
- monostylism, 17, 37
- morphology
 - overview of characteristics, 202-5
 - summary, 202-45
 - see also* nominal morphology; verbal morphology
- morphosyntax, 17, 21, 32, 102-49, 174, 175, 177, 179, 180, 181, 182, 184, 214, 243, 244, 250, 258, 259
 - base features, 19, 20
 - complex, 25
 - conservative, 19
 - decay of native, 51
 - effects of language decline on, 36-7
 - see also* nominal morphosyntax; verbal morphosyntax
- Mull, 40, 44, 163
- Multi-dimensional (MD) analysis, 18, 32-4, 38, 53, 63, 191-3
- Munster, 36, 40
- mutation, initial *see* lenition

N

- narrative
 - Biber's dimensions of, 33, 34, 192
 - compared with non-narrative discourse, 148, 187
 - register marking, 33, 34
 - single person, 25
 - spoken compared with written, 26, 28
 - see also* fiction; short stories; story-tellers; traditional narrative
- nasalisation, 198, 208

- nativisation of English loans, 68
- negation, 238, 241
 - with *ann*, 247-8
 - clausal, 65, 111-12, 148
 - conditionals, 261
 - double negative constructions, 241
 - with *gun*, 241
 - imperatives, 241
 - prefixes, 241
- Nelson, Gerald, 18, 27, 28, 30, 38, 52, 63, 74, 76, 84, 86, 94, 177, 300, 301, 304
- news scripts, 17, 21, 52, 182
 - complement clauses in, 94
 - lack of discourse freedom in, 76-7
 - low prevalence of past tense in, 106
 - profile of register, 94, 182-3, 187
 - situational description and examples, 62-3
 - temporal reference in, 92
 - texts for study, 55
- news-speak, 17-18, 48, 62
- newspapers, 49, 62
- nexus, 252, 254
 - see also* coordination; cosubordination; subordination
- Ní Laoire, Siobhán, 19, 20, 35, 36, 145, 304
- nominal categorisers, 228
- nominal concordance, 151-2, 186
- nominal morphology, 65, 205-28
- nominal morphosyntax, 116-28
- nominalisation, 26, 228
- nominals, predicate, 245-8
- nominative case, 209
 - concordance data, 151-2
- non-Gaelic speakers, effect of presence, 46, 47
- nonconcordance
 - case-marking, 150-7
 - definition, 125
 - frequency per tag type, 156
 - grammatical, 65, 125-6
- noun clauses, 25
- noun phrases, 30
 - complex *see* complex NPs (noun phrases)
 - complexity and function, 157-71, 195
 - density, 32
 - extraclausal, 79, 98, 158
 - grammar and complexity, 150-71, 195
 - NP-Clause structures, 29
 - subordination, 262-3
- nouns
 - alienable vs. inalienable, 206
 - as anaphors, 206

- assimilated and unassimilated borrowings, 66
- borrowed from English, 143-4
- case, 209-12
- classes, 206-8
- common, 205
- complex, 157, 186
- with defective gender, 68
- definiteness/referentiality, 208-9
- dialectal variation in, 68
- gender, 206
- genitive, 65, 126-7
- indefinite marking, 65, 127-8, 148
- mass, 206
- morphology, 205-13
- number marking, 206
- plurals, 212
- proper, 205
- proportion of noncordant to overall, 65, 125-6, 154, 155
- total, 65, 123-4
- see also* verbal nouns
- nucleus, 12, 31, 67, 252, 253
- Nukulaelae Tuvaluan, 20, 32
- numerals, 66, 145-6, 218-19
 - cardinal, 218-19
 - ordinal, 219
 - pronouns, 219
 - use of English for large, 219

O

- Ó Maolalaigh, Roibeart, 35, 201, 217, 304
- objects
 - genitive, 26
 - postposition of weak pronominal, 249-50
- obsolescence, 17
- occupation, 55
 - and Gaelic-English usage, 45, 46
- Ochs, Eleanor, 17, 24, 25, 26, 28, 38, 92, 94, 116, 190-1, 193, 301, 304
- O'Donnell, Roy C., 25
- official documents, 33, 50
- Oftedal, Magne, 198, 207, 304
- operators, 253
- oral culture, transmission of, 35
- oral folk tale sample text, 275-8
- oral language *see* spoken language
- oral literature, 19, 24, 27, 35, 36, 50, 298
 - high register greeting language in, 36
- oral-literate differences, 28-9, 34
 - clefts, 74
 - information amount, 150
 - see also* speech and writing

orality, 26, 179, 272
 origins, marking by speech style, 36
 orthography, 49, 55, 143, 182, 198, 201, 208, 213
 Outer Hebrides, 42, 43, 45, 201, 208, 297, 298
 dialects, 40
 economics of, 43
 sonorants, 199

P

palatalisation, 152, 198, 200-1, 204, 205, 206, 207, 210, 211, 212, 221
 paralinguistic strategies, 26, 28, 29
 see also gestures
 parataxis, 30
 parenthetical statements, 271-2
 parliamentary terms, 48
 participants
 age, 22
 attitudinal stance, 34, 57
 characteristics, 55
 literacy levels, 25
 relationships, 20
 participles, 25, 26, 28, 29, 122, 244-5, 254, 263
 passive, 242-3
 particles
 discourse, 271
 pre-verbal, 230
 parts of speech, frequencies of different, 25
 passive clauses
 with agents, 65, 108-9
 frequency, 65, 109-10
 passive voice, impersonal forms, 280
 passives, 24, 25, 26, 72, 148, 174, 178, 179, 181, 183, 186, 187, 191, 192, 280
 construction types, 242-4
 definition, 242
 impersonal, 242, 244
 inflected or synthetic, 108, 110
 move from synthetic to analytic, 18
 participle, 242-3
 periphrastic or analytic, 108
 personal, 242
 valence decreasing expressions, 65, 107-11, 242-4
 passivisation, 34
 past tense, 26, 33, 34, 106-7
 definite, 65
 indefinite (past-habitual), 65, 107, 235, 236
 indefinite (present-future), 65, 107, 235-6
 progressive, 228
 simple, 228
 past-conditional copula, 245, 246

- past-habitual, 65, 107, 235, 236
- pauses, 29, 55
- Payne, T.E., 103, 206, 221, 237, 242, 243, 248, 304
- pedagogy, 19
- perfect aspect, 65, 102-3, 237
- periodicals, 49, 50
- periphrasis, preferred over inflection, 274
- persuasion, 66
 - marking of, 33, 34
- Perthshire, East, 201
- phonetics, 19, 23, 35, 273
- phonology, 19, 32, 42, 68, 143, 273, 302
 - summary, 198-201
- phrases, better unit of analysis of speech than clauses, 30
- pitch, 29, 97, 130
- placement
 - fronted adverbials and peripheral elements, 64
 - postposed subordinate clause, 64
 - preposed subordinate clause, 64
- plannedness, 17, 28, 50, 188, 190
- pluperfect aspect, 237
- poetry, Gaelic, 49, 50
- Poole, M.E., 25, 28, 190, 304
- Popular Tales of the West Highlands*, 23, 298
- popular writing, 21, 52
 - definition, 54
 - profile of register, 184-5
 - situational description and examples, 62
 - texts for study, 54
- possession, 128-31
 - with *aig*, 66, 128, 129, 130-1, 213
 - alienable vs. inalienable, 212
 - pronominal, 66, 128, 213, 227
 - simple, 129-30
 - summary, 212-13
- possessive predication, 246-7
- pragmatics of reference, 29
- preaspiration, 11, 40, 201, 304
- predicate adjectives, 246
- predicate nominals, 245-8
 - proper inclusion and equation, 245-6
- prefixes, 203
 - negative, 241
- prepositions, 66, 131-2
 - complex, 227-8, 249
 - compound, 227
- prounouns, 226
 - simple, 224-7
 - summary, 224-7
 - total of, 66

- prescriptivism, 19, 24, 195
- present perfect aspect, 24
- present tense, 25, 33, 65, 105-6
 - historical, 26, 107
- 'present-future' tense, 65, 235-6, 245
- preverbal constituents, 25
- primacy orientation, 23-4, 37
- producer characteristics, 189-90
- production constraints, 25, 34, 71, 148, 149, 178, 188-9, 190, 193
- profanity, 36
- proficiency
 - and age, 46
 - differing levels, 37
- profiles of individual registers, 172-85, 186
- progressive aspect, 25, 65, 103-5, 228, 237
- pronominal objects, postposition of weak, 249-50
- pronouns, 66, 132-5
 - demonstrative, 66, 132, 133, 215
 - emphatic suffix, 214
 - first person, 26
 - honorifics, 135, 214
 - indefinite, 217-18
 - interrogative, 216-17
 - numerical, 219
 - personal, 132, 214
 - possessive, 66, 128, 213
 - prepositional, 214, 226
 - prepositional possessive, 214, 227
 - reciprocals, 216
 - reflexives, 215-16
 - relative, 217
 - second person, 214
 - summary, 214-18
 - third person, 33, 34
 - total of, 66, 133-4
 - in traditional narrative, 134
- prose
 - planned written, 17, 50
 - see also* academic prose; formal prose
- prospective aspect, 65, 105, 237
- pseudo-clefts, 269-70
- public speaking, 36
- publishing, Gaelic, 49-50
- purity, 48
- purpose, 31, 33

Q

- quantifiers, 26
- quantitative methods, 18

- question words, 216-17
 - form compared with relative form, 217
- questions
 - indirect, 257
 - polar (yes/no), 239-40
 - tag, 240
 - wh*-, 240

R

- radio, 48, 51, 54, 58
 - see also* news scripts; news-speak
- radio interviews, 21, 52, 187, 194
 - example of tagged, 67
 - profile of register, 175-6
 - situational description and examples, 56-7
 - texts for study, 54
- Ramchand, Gillian, 19, 249, 305
- Rathlin Island, 40
- reading ability, 49
- reduced clauses *see* small clauses
- reference
 - explicit vs. situation-dependent, 192
 - extra-textual, 120, 136
 - intra-textual, 84, 91, 92, 119, 136
- referential devices, 264-5
- register
 - definitions, 19, 22, 31, 55
 - formal, 19
 - informal, 19
 - and language choice, 47-8
 - and mode conflated, 28
 - problems with definition, 27-8
 - use of term, 31, 186
 - see also* spoken registers; written registers
- register markers, 32, 35
- register range, 47-50
- register studies, 31-2
 - assumptions and limitations, 22, 31-2
 - macroscopic analysis, 33
 - methodology, 32, 52-72
 - microscopic analysis, 33
 - multi-dimensional models, 18, 32-4, 38, 53, 63, 191-3
 - obstacles, 20, 32
 - past research, 190-4
 - previous Scottish Gaelic, 35-7
 - profiles of individual registers, 172-85, 186
 - research questions, 21
 - single register description, 32
 - speech and writing, 23-38

- register variation, 18, 19, 20, 21, 22, 32, 35, 71, 72, 87, 96, 102, 103, 105, 107, 145, 148, 150, 300, 301
 - Biber's dimensions of, 33-4, 191-3
 - Chafe's dimensions, 26, 28, 30, 37, 191, 194
 - English, 17
 - persistence in language decline, 37
 - research findings, 36
 - universal dimensions, 33-4, 38, 193, 195, 196
- relative clauses, 28, 29, 64, 68, 92-3, 100, 159, 262-3
 - complexity in formal prose, 164
 - cosubordination, 264
 - non-restrictive, 29
 - restrictive, 29
 - single, 164-6
- relativisation strategies, 263
- religion, 36, 42, 43
- religious language, 19, 35, 36, 47
 - see also* Biblical language
- repetition, 26
- rhetorical mode, use of term, 31
- rhetorical units, sentences as, 30
- rhythm, 29
- ritual language, 26-7, 59
- Role and Reference Grammar (RRG), 29, 251-4, 264
- Romaine, S., 32, 52, 305
- Russian, 20, 29, 48, 158, 247

S

- Sapir, Edward, 23, 305
- satire, 36
- School of Scottish Studies, 35, 54, 59
- scope, 77, 174, 176, 178, 182
 - definition of, 76
 - see also* discourse freedom
- Scottish English, 29, 52
- Scottish Gaelic Corpus
 - data analysis, 70
 - data collection and composition, 52-63
 - mark-up, 66-9
 - methodology, 52-72
 - participants in, 55
 - proposals for large, 196
 - selection of linguistic features, 63-6
 - tags, 64-9
 - texts listed, 287-90
- seanchaidhean*, 35, 59
- self-consciousness, 149, 194
- semi-speakers, 44, 47
- Seneca, 26

sentences

- ill-suited for analysis of informal speech, 30, 194

- length of, 18, 24, 25, 30

- one-word, 27

- Shapiro-Wilk test, 70

- Shaw, John, 36

- shibboleth, 42

- short stories, 49, 50, 54

- mimic spontaneous speech, 26

- short-term memory, limits of, 29, 186, 191

- simplification, 150, 209, 274

- Skye, 42, 45

- small clauses, 242, 250-1

- social context, 20

- sociolinguistics, 30, 35, 39-51

- Somali, 32

- speech and writing, 23-4

- Biber's dimensions, 33-4

- complexity and subordination in, 28-9, 37

- linguistic differences between, 24-8

- spoken language

- beul* vs. *cainnt*, 36

- characteristics, 25

- complexity, 25

- divergence from written language, 18-19, 23-4, 26, 27

- influence of English on, 273-5

- interference from written forms, 36

- as syntactically richer, 28

- spoken registers, 47-9

- sub-corpus, 21, 52

- Spontaneous Spoken Language (SSL), 21, 29-31, 38, 194

- adult compared with early child discourse, 25

- characteristics, 25, 29-31

- compared with formal prose for complexity and function, 150-71

- corpus needed, 196

- observations, 194-5

- properties, 194-5

- syntax and morphology, 30, 150-71

- see also* colloquial speech; conversation

- sports broadcast, 21, 27, 52

- situational description and examples, 57-9

- texts for study, 54

- sports reportage, 187, 192, 194

- profile of register, 176-8

- statistics

- descriptive, 32

- inferential, 18, 25, 32

- multivariate, 32-3, 37

- used in data analysis, 70

- status, 17, 114, 135
 - in language decline, 37
 - replacement of low by higher, 17
- stem modification, 204
- story-tellers, 35, 59
- strong verbs, 229, 234
- structured interviews, 25
- style
 - marking of people's origins by, 36, 42
 - use of term, 22, 31
- subjects
 - dative, 251
 - genitive, 26
- subordinate clauses, 18, 26, 90-7, 253
 - cause or reason, 64, 96-7
 - distribution and types, 27, 28, 29
 - finite, 28
 - fronted in traditional narrative, 179
 - non-finite, 28, 253
 - placement of, 82, 83-4, 100
 - postposed, 64
 - preposed, 25, 64
 - temporal, 64, 90-2, 101, 262
 - see also* complement clauses; concessive clauses; conditional clauses; relative clauses
- subordination, 17, 26, 27, 33, 37, 68, 71, 72, 73, 82, 84, 87, 90, 91, 92, 96, 97, 100, 101, 175, 177, 179, 182, 187, 191, 193, 250, 251, 252, 256-63, 297, 301, 302, 303
 - adverbial clauses, 257-62
 - complement clauses, 256-7
 - juxtaposition versus, 97-100, 101, 195
 - morphosyntax of, 258-9
 - noun phrase, 262-3
 - prevalence of, 25, 29, 195
 - problems with the concept, 28-9
 - reason, 258-9
 - in speech and writing, 28-9
 - temporal, 262
- suffixes, 203-4
- superstratum influences, 37
- suppletion, 204, 221
- suppletive (irregular) verbs, 279
- Survey of the Gaelic Dialects of Scotland*, 35, 42
- svarabhakti, 40
- Sweet, Henry, 23
- syndesis, 26
- syntactic reduction, 35, 36
- syntax
 - functional, 19
 - manipulation of structure for stylistic variation, 20
 - summary, 245-75
- synthesis, 202

T

t-units, 25

tag set

definition, 22

full, 64-6, 196

full listed, 291-5

tagging, 18, 32, 63, 66, 68, 196

tags

clause-level, 67

counting, 69

key to, 67

word-level, 67

Taiwanese, 20, 32, 34, 62, 114, 115, 192, 301

TAM system (tense-aspect-mode), 19, 102, 233-41

Tannen, Deborah, 26, 28, 29, 106, 107, 127, 191, 297, 298, 305

tape recordings, 20, 35

technical discourse, 47

television, effect on Gaelic-English usage, 46, 48, 49, 51, 142, 213, 287

tense, 32, 105-7

definite mode, 234

indefinite, 1, 65, 107, 234

Macaulay's paradigm, 234-6

narrative compared with non-narrative discourse, 148

summary, 233-6

see also future tense; past tense; present tense

Ternes, E., 198, 199, 305

text-types, 20, 31

comparison of Jang and Biber, 34

textbooks, Gaelic-medium, 47-8, 50

texts

internal or external criteria for grouping, 31

origin and selection of, 53-63

Thompson, Derick, 47

Thurneysen, Rudolph, 202

time factors/constraints, 28, 29, 62

time references, 187

tourism, 43

traditional narrative, 21, 36, 52, 114, 148, 194

conservative grammar in, 111

profile of register, 179-80

situational description and examples, 59-60

texts for study, 54

transcription, 16, 23, 35, 53, 55, 68, 292

of speech as arbitrary, 30

translation, 48

transmission

intergenerational, 44, 50

second-generation, 42

triphthongs, 199

type-token ratio, 25, 66, 146, 147-8, 176, 178, 181, 182, 185, 191, 195
 definition, 25

U

Uists

dialect, 22, 40, 42, 53, 54, 57, 60, 80, 199, 207, 208, 226, 235, 246, 260, 267, 272, 273,
 275, 279, 287, 298, 303
 Gaelic-speaking adults, 46
 Gaelic-speaking children, 44
 map, 41

universal dimensions of register, 33-4, 38, 193, 195, 196

urban Gaels, 42

V

valence decreasing expressions, 65, 107-11

passives and impersonals, 242-4

valence increasing expressions, 241-2

Van Valin, Robert, 29, 68, 238, 251, 252, 253, 254, 270, 300, 306

variation

context-based linguistic, 31-2

contextual and functional, 18, 21, 32

dialectal, 22, 199, 201, 206, 208, 210, 213, 214, 215, 224, 230, 232, 246, 259, 266, 270,
 279, 280, 4068

see also register variation

verb classes, 231-2

verbal categorisers, 241-4

verbal modifiers, 244-5

verbal morphology, 228-45

verbal morphosyntax, 65, 102-16

verbal nouns, 132, 206, 244-5

structure and distribution, 230-1

verbals, non-finite, 25

verbs

composite, 251

consonant-initial, 235

distribution, 229

dynamic or stative, 231

fusion, 202

impersonal affixes, 233

passive, 25

personal affixes in finite and non-finite, 232-3

pre-verbal particles, 230

private, 66, 140, 141-2, 149

public, 33, 66, 140, 142

structure, 229-30

suasive, 66, 141, 142-3

summary, 229-41

synthetic older forms, 36

types, 140-3

- vowel-initial, 235
- see also* copular verbs; strong verbs; suppletive (irregular) verbs; weak verbs
- vocabulary, new terms, 47-8
- voice quality, 29
- vowel harmony, 211
- vowels, 198

W

- weak verbs, 229, 234
 - inflection, 234-6
- Weinert, Regina, 29, 90
- Weitzman, Michael, 18, 27, 28, 38, 94, 301
- Western Isles Language Plan (2005), 50
- Wilson, Andrew, 31, 42, 196, 303
- word length, 17, 24, 33, 66, 146-7, 187, 191, 192
- word order, 20
 - highlighting of constituents, 266-71
 - non-canonical, 270
 - VSO, 228, 229
- word stress, 199
- words
 - 100 most common, 281-2
 - index of those used in grammar, 283-6
- writers, Gaelic, 49-50, 51
- writing
 - density of, 25, 28
 - interference from spoken forms, 36
 - see also* popular writing; speech and writing
- written language
 - characteristics, 25
 - divergence from spoken language, 18-19, 23-4, 26
 - emphasis on, 23-4, 37
- written registers, 49-50
 - sub-corpus, 21, 52

Y

- Yoruba, 27

This is the first in-depth, quantitative study of register variation in Scottish Gaelic, an endangered Celtic language. Previous work on the subject has been mainly anecdotal in nature or a by-product of other lines of investigation. The study confirms that it is impossible to uncover the characteristics of any register in Gaelic fully without benchmark data on the language's overall register range and morpho-syntactic variation, now provided by the present ground-breaking study.

The book shows robust differences between most of the register types in Scottish Gaelic. It also shows major contrasts between narrative and non-narrative registers, reportage and non-reportage registers, and formal writing and other register types. It identifies five underlying contextual parameters to be responsible for the variation between registers in Scottish Gaelic: production constraints, discourse freedom, information orientation, interaction, and producer characteristics.

Overall, the book's results correlate well with those of the current leading, most influential models of register and text-type variation. The book demonstrates that Scottish Gaelic, despite being an endangered language, has a level of register variation comparable to that found in larger languages such as English, with whom the patterns of common variation are universal.

The book also presents a new morpho-syntactic grammar of contemporary spoken and written Scottish Gaelic, providing the fullest and most authoritative linguistic description currently available.

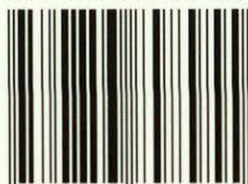
"This book is accessible to Gaelic speakers as it is to descriptive linguists. It provides compulsory reading for anyone interested in knowing about the morphology, syntax and discourse features of contemporary Scottish Gaelic." (Dr John M. Kirk)

"This book has much to offer all linguists concerned with the grammar of spoken language." (Prof. Jim Miller)

Dr William Lamb is a lecturer in Scottish Gaelic and Music at Lews Castle College, Benbecula (UHI Millennium Institute) and lives on the Isle of North Uist. Originally from Baltimore, Maryland, William Lamb came to Scotland in 1996 to take an MSc in Scottish Gaelic Studies at the University of Edinburgh, where, in 2002, he graduated with a PhD on the subject of the present book.



ISBN 978-0-85389-895-5



9 780853 898955 >